

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 11:56:10 ; Search time 113 Seconds

(without alignments)

1419.044 Million cell updates/sec

Title: US-09-879-248-6

Perfect score: 2310

Sequence: 1 MSILTLNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

complete

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2310	100.0	447	2	AAW96260 Hypersens
2	2310	100.0	447	3	AAW71094 Erwinia a
3	2310	100.0	447	3	AAW94855 A hyperse
4	2310	100.0	447	5	AAO22548 Hypersens
5	2310	100.0	447	5	AAE18296 Erwinia a
6	2310	100.0	447	5	AAE16448
7	2310	100.0	447	5	ABBO9226
8	559	24.2	424	2	AAW97851
9	559	24.2	424	3	AAW71098
10	559	24.2	424	3	AAW94859 A hyperse
11	559	24.2	424	5	AAO22550 Hypersens
12	559	24.2	424	5	AAE18298
13	559	24.2	424	5	AAE16452
14	559	24.2	424	5	ABBO9228
15	359.5	15.6	197	2	AAW83014 Bacillus
16	359.5	15.6	197	2	AAW77412 Bacillus
17	341	14.8	221	2	AAW28446
18	341	14.8	221	8	ADK48880 Bacillus
19	199.5	8.6	1306	6	ABU36481 Protein e
20	193	8.4	694	6	ABU36802 Protein e
21	191.5	8.3	594	6	ABU22680 Protein e
22	191	8.3	280	4	ABBE1650 Drosophil
23	188	8.1	639	6	ABU36685 Protein e
24	186.5	8.1	591	6	ABU36439 Protein e
25	185.5	8.0	1381	6	ABU36971 Protein e

26	185	8.0	588	7	ADM40790	Adm40790 Mycobacte
27	183	7.9	588	6	ABU36945	Abu36945 Protein e
28	182	7.9	606	6	ABU36540	Abu36540 Protein e
29	181.5	7.9	484	6	ABU15873	Abu15873 Mycobacte
30	181.5	7.9	484	6	ABU36943	Abu36943 Protein e
31	181.5	7.9	484	7	ABO23518	Abu23518 Mycobacte
32	181.5	7.9	2411	3	ABO23860	Abu23860 Haemophil
33	180.5	7.8	1079	7	ABO23519	Abu23519 Mycobacte
34	179	7.7	730	3	AAO29582	Aag29582 Arabidops
35	179	7.7	752	3	AAO29581	Aag29581 Arabidops
36	176.5	7.6	562	6	ABU36590	Abu36590 Protein e
37	176.5	7.6	2090	6	ABBE4682	Abb4682 Drosophil
38	174	7.5	615	6	ABU36862	Abu36862 Protein e
39	174	7.5	2042	2	AAW56319	Aaw56319 Haemophil
40	174	7.5	2586	4	ABBE6878	Abb6878 Drosophil
41	172.5	7.5	584	6	ABU36979	Abu36979 Protein e
42	171.5	7.4	667	6	ABU36570	Abu36570 Protein e
43	171	7.4	505	6	ABU34417	Abu34417 Protein e
44	169.5	7.3	914	6	ABU36682	Abu36682 Protein e
45	169.5	7.3	525	6	ABU36839	Abu36839 Protein e
46	168	7.3	1357	4	ABB70775	Abb70775 Drosophil
47	167.5	7.3	1011	6	ABU36665	Abu36665 Protein e
48	166.5	7.2	518	6	ABU22710	Abu22710 Protein e
49	166.5	7.2	532	6	ABU36734	Abu36734 Protein e
50	166	7.2	461	6	ABU36826	Abu36826 Protein e
51	166	7.2	603	6	ABU36613	Abu36613 Protein e
52	166	7.2	767	6	ABU20460	Abu20460 Protein e
53	165.5	7.2	344	2	AAW61116	Aaw61116 Hypersens
54	165.5	7.2	344	2	AAW62457	Aaw62457 Pseudomon
55	165.5	7.2	344	2	AAW75865	Aaw75865 Pseudomon
56	165.5	7.2	344	2	AAW87641	Aaw87641 A hyperse
57	165.5	7.2	344	2	AAW71099	Aay71099 Pseudomon
58	165.5	7.2	344	3	AAW84860	Aay84860 A hyperse
59	165.5	7.2	344	3	AAW84860	Aay84860 A hyperse
60	165.5	7.2	344	4	AAO26712	Aao26712 Pseudomon
61	165.5	7.2	344	5	AAO22551	Aao22551 Hypersens
62	165.5	7.2	344	5	AAE18299	Aae18299 Pseudomon
63	165.5	7.2	344	5	AAE16453	Aae16453 P. solana
64	165.5	7.2	344	5	ABBO9229	Abe9229 P. solana
65	164.5	7.1	898	2	AAW31853	Aaw31853 Mycobacte
66	163	7.1	618	6	ABU36684	Abu36684 Protein e
67	161.5	7.0	576	6	ABU36623	Abu36623 Protein e
68	161.5	7.0	937	6	ABU36438	Abu36438 Protein e
69	161	7.0	1600	4	ABBE4927	Abb4927 Drosophil
70	161	7.0	1601	8	ADO00994	Ado00994 Fruit fly
71	161	7.0	1601	8	ADO07915	Ado07915 Fly polyp
72	160.5	6.9	957	6	ABU36437	Abu36437 Protein e
73	160	6.9	1096	8	ADP99004	Adp99004 C. albica
74	159.5	6.9	778	6	ABU36827	Abu36827 Protein e
75	159	6.9	714	5	ABP73702	Abp73702 Candida a
76	159	6.9	3286	6	ABU21773	Abu21773 Protein e
77	158	6.8	700	4	ABBE60376	Abb60376 Drosophil
78	157.5	6.8	2870	3	AAV95559	Aay95559 Caenorhab
79	157.5	6.8	2870	8	ADNO7634	Adn07634 Caenorhab
80	157.5	6.8	3178	3	AAV95556	Aay95556 Caenorhab
81	157.5	6.8	3178	8	ADNO7623	Adn07623 Caenorhab
82	157	6.8	491	6	ABU19448	Abu19448 Protein e
83	157	6.8	979	6	ABU19448	Abu19448 Protein e
84	156	6.8	2039	2	AAW56322	Aaw56322 Haemophil
85	155.5	6.7	2353	2	AAW99393	Aar99393 Haemophil
86	155	6.7	385	2	AAW45751	Aar45751 Erwinia a
87	155	6.7	495	2	AAW39747	Aaw39747 Human mar
88	155	6.7	1113	6	ABR53293	Abr53293 Protein e
89	155	6.7	1113	7	ADK63560	Adk63560 Disease t
90	155	6.7	1672	6	ABU27330	Abu27330 Protein e
91	154.5	6.7	354	7	ABO23520	Abu23520 Mycobacte
92	154.5	6.7	630	6	ABU26539	Abj26539 Aspergill
93	154.5	6.7	975	6	ADA33477	Ada33477 Acinetoba
94	154	6.7	334	6	ABU20015	Abu20015 Protein e
95	154	6.7	571	5	ABP39020	Abp39020 Staphyloc
96	154	6.7	651	6	AAE36884	Aae36884 Argiope t
97	154	6.7	657	6	ABBS7783	Abbs7783 Drosophil
98	154	6.7	1099	6	ABU20296	Abu20296 Protein e

99 154 6.7 3275 4 ABB70437 Abb70437 Drosophil
100 154 6.7 3275 7 ADK11301 Adk11301 Drosophil

ALIGNMENTS

RESULT 1
AAW96260
ID AAW96260 standard; protein; 447 AA.
XX
AC AAW96260;
XX
DT 14-JUN-1999 (first entry)
XX
DE Hypersensitive response eliciting protein (HrpW).
XX
KW Hypersensitive response elicitor protein; hypersensitive response; hrpW;
KW pathogen; infection; crop protection; disease resistance;
XX pest resistance; transgenic plant; colouration; maturation.
XX
OS Erwinia amylovora.
XX
XX WO9907208-A1.
XX
PD 18-FEB-1999.
XX
PF 27-JUL-1998; 98WO-US015547.
XX
PR 06-AUG-1997; 97US-0055108P.
XX
PA (CORR) CORNELL RES FOUND INC.
XX
PI Kim JF, Beer SV;
XX
DR WPI; 1999-167126/14.
DR N-PSDB; AAX09007.
XX
PT New Erwinia amylovora hypersensitive response eliciting gene and protein
PT - useful for providing transgenic plants and seeds with enhanced growth,
PT and insect and disease resistance.
XX
PS Claim 1; Page 50-51; 54pp; English.
XX
CC The hypersensitive response eliciting protein (hrp) or polypeptide is
CC produced as part of an active defense by plants against incompatible
CC pathogen infections. The hypersensitive response is a rapid localised
CC necrosis. The hrp protein and gene when used in nucleotide constructs are
CC useful for providing disease resistance to plants, insect control to
CC plants, and enhancing plant growth (enhancing fruit size and earlier
CC colouration and maturation), by direct application of the protein to
CC plants, or by producing transgenic plants or seeds using the hrp gene
XX
SQ Sequence 447 AA;

Query Match 100.0%; Score 2310; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQOPIDQRTIEQMAQLLAELLSLIS 60
DB 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQOPIDQRTIEQMAQLLAELLSLIS 60

QY 61 POSGNAATGAGNDQTTGVGNAGLNGRGTAGTTTQSDSQNMLSEMNGNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGVGNAGLNGRGTAGTTTQSDSQNMLSEMNGNGLDQAITPDG 120

QY 121 QGGGQIGDNPLLKAMUKLTAARMWDGSDQFGQPGTGNNASAGTSSGSGSPNDLSGGKA 180
DB 121 QGGGQIGDNPLLKAMUKLTAARMWDGSDQFGQPGTGNNASAGTSSGSGSPNDLSGGKA 180

QY 181 PSGNSPSGNYSPVSTFSPPTSTPTSPLDPPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240
DB 181 PSGNSPSGNYSPVSTFSPPTSTPTSPLDPPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240

DB 181 PSGNSPSGNYSPVSTFSPPTSTPTSPLDPPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKQOTFTAGSELGDDGQSQENQKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKQOTFTAGSELGDDGQSQENQKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGIHLYGDADKIDNLHVTNVGDAITVKPNSACKKSHVEITNSFFHASD 360
DB 301 LKNVTMGDDGADGIHLYGDADKIDNLHVTNVGDAITVKPNSACKKSHVEITNSFFHASD 360
QY 361 KILQLNADTNLSVDNKKAKDFGTFVRTNGQQQGNWDNLNLSHISAEDCKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNKKAKDFGTFVRTNGQQQGNWDNLNLSHISAEDCKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 2
AAY71094
ID AAY71094 standard; protein; 447 AA.
XX
AC AAY71094;
XX
DT 08-SEP-2000 (first entry)
XX
DE Erwinia amylovora hypersensitive response elicitor #2.
XX
KW Hypersensitive response elicitor; environmental stress resistance; plant.
XX
OS Erwinia amylovora.
XX
PN WO200028055-A2.
XX
PD 18-MAY-2000.
XX
PF 04-NOV-1999; 99WO-US026039.
XX
PR 05-NOV-1998; 98US-0107243P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Schading RL;
XX
DR WPI; 2000-376566/32.
DR N-PSDB; AAD00669.
XX
PT Application of a hypersensitive response elicitor protein to plants to
PT impart stress resistance.
XX
PS Disclosure; Page 10-12; 84pp; English.
XX
CC The patent discloses a method to impart stress resistance to plants by
CC applying a hypersensitive response elicitor in a non-infectious form to a
CC plant or seed. The present sequence is a hypersensitive response elicitor
CC protein from Erwinia amylovora. The protein is heat stable, protease
CC sensitive and suppressed by inhibitors of plant metabolism. It is used to
CC impart stress resistance to plants
XX
SQ Sequence 447 AA;

Query Match 100.0%; Score 2310; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQOPIDQRTIEQMAQLLAELLSLIS 60
DB 1 MSILTLNNNTSSPGLFQSGGDNGLGHNANSALGQOPIDQRTIEQMAQLLAELLSLIS 60

QY 61 POSGNAATGAGNDQTTGVGNAGLNGRGTAGTTTQSDSQNMLSEMNGNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGVGNAGLNGRGTAGTTTQSDSQNMLSEMNGNGLDQAITPDG 120

QY 121 QGGGQIGDNPFLKAMLKLIARMMDGSDQFGQGTGNNSSAGTSSGGSPFNDSGGKA 180
 DB 121 QGGGQIGDNPFLKAMLKLIARMMDGSDQFGQGTGNNSSAGTSSGGSPFNDSGGKA 180
 QY 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAAGGSTPVTDHDPVGSAGIGAG 240
 DB 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAAGGSTPVTDHDPVGSAGIGAG 240
 QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQGSQENQKPLFIEDGAS 300
 DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQGSQENQKPLFIEDGAS 300
 QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFHASD 360
 DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFHASD 360
 QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDEGLN 420
 DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDEGLN 420
 QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
 DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 3
 AAY84855
 ID AAY84855 standard; protein; 447 AA.
 AC AAY84855;
 XX
 DT 08-AUG-2000 (first entry)
 DE A hypersensitive response elicitor protein.
 XX
 KW Hypersensitive response; insect control; disease resistance;
 KW Hypersensitive response elicitor; plant growth; vegetable; crop;
 KW ornamental plant.
 XX
 OS Erwinia amylovora.
 XX
 PN WO200020452-A2.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-US023181.
 XX
 PR 05-OCT-1998; 98US-0103050P.
 XX
 PA (EDEN-) EDEN BIOSCIENCE CORP.
 XX
 PI Wei Z, Fan H, Niggemeyer JL;
 XX
 DR WPI; 2000-303745/26.
 DR N-PSDB; AAA14939.
 XX
 PT Hypersensitive response elicitor polypeptides useful for imparting
 PT enhanced growth, disease resistance and insect resistance to plants,
 PT especially vegetables and ornamental flowers.
 XX
 PS Disclosure; Page 12-13; 100pp; English.
 XX
 CC The present sequence represents a hypersensitive response elicitor
 CC polypeptide. The specification describes hypersensitive response elicitor
 CC polypeptide fragments, which do not elicit a hypersensitive response.
 CC Instead, the proteins impart disease resistance to plants, enhance plant
 CC growth, and/or control insects. The polypeptide fragments may be used to
 CC these properties to plants. The plants which may be treated in this way
 CC include vegetables, crops and ornamental plants such as alfalfa, rice,
 CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet
 CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,
 CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,

CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,
 CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,
 CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis
 CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,
 CC carnation or zinnia
 XX
 SQ Sequence 447 AA;
 Query Match 100.0%; Score 2310; DB 3; Length 447;
 Best Local Similarity 100.0%; Pred. No. 2.4e-151;
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTLTNNNTSSPGLFQSGDNGLGGHNANSALGOOPIDRQTIQMAQLLAEKLSLLS 60
 DB 1 MSTLTNNNTSSPGLFQSGDNGLGGHNANSALGOOPIDRQTIQMAQLLAEKLSLLS 60
 QY 61 POSGNAATCAGGNDQTTGVGNAGLNGRGTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120
 DB 61 POSGNAATCAGGNDQTTGVGNAGLNGRGTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120
 QY 121 QGGGQIGDNPFLKAMLKLIARMMDGSDQFGQGTGNNSSAGTSSGGSPFNDSGGKA 180
 DB 121 QGGGQIGDNPFLKAMLKLIARMMDGSDQFGQGTGNNSSAGTSSGGSPFNDSGGKA 180
 QY 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAAGGSTPVTDHDPVGSAGIGAG 240
 DB 181 PSNGSPSGNYSVSTSPSTPTSPPLDPSSPTKAAGGSTPVTDHDPVGSAGIGAG 240
 QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQGSQENQKPLFIEDGAS 300
 DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQFTTAGSELGQGSQENQKPLFIEDGAS 300
 QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFHASD 360
 DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFHASD 360
 QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDEGLN 420
 DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFVKSDEGLN 420
 QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
 DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 4
 AAO22548
 ID AAO22548 standard; protein; 447 AA.
 XX
 AC AAO22548;
 XX
 DT 28-OCT-2002 (first entry)
 XX
 DE Hypersensitive response elicitor protein #2 from Erwinia amylovora.
 XX
 KW Desiccation inhibitor; longevity enhancer; desiccation inhibitor;
 KW cutting; ornamental plant; hypersensitive response elicitor protein;
 KW transgenic ornamental plant; early flowering; longevity enhancer; plant;
 KW flower bloom; flower.
 XX
 OS Erwinia amylovora.
 XX
 PN WO200237960-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 06-NOV-2001; 2001WO-US043715.
 XX
 PR 13-NOV-2000; 2000US-0248169P.
 XX
 PA (EDEN-) EDEN BIOSCIENCE CORP.
 XX
 PI Wei Z, Leon E, Oviedo A;

```
XX WPI; 2002-575194/61.
DR N-PSDB; AAL41133.
XX
PT Inhibiting desiccation of cuttings from ornamental plants, by treating
PT ornamental plants with hypersensitive response elicitor protein, or
PT expressing heterologous hypersensitive response elicitor protein in
PT plants.
XX
PS Disclosure; Page 12-13; 69pp; English.
XX
CC The invention relates to a method for inhibiting desiccation of cuttings
CC from ornamental plants. The method involves treating the cuttings with a
CC hypersensitive response elicitor protein or polypeptide, or providing a
CC transgenic ornamental plant or plant seed transformed with a DNA molecule
CC encoding the hypersensitive response elicitor polypeptide, and growing
CC the ornamental plant or transgenic ornamental plant produced from the
CC transgenic ornamental plant seed. The hypersensitive response elicitor
CC protein or polypeptide is useful for inhibiting desiccation of cuttings
CC from ornamental plants, for harvesting cuttings from ornamental plants,
CC for promoting early flowering of ornamental plants, and enhancing the
CC longevity of flower blooms on ornamental plant cuttings. This sequence
CC represents a hypersensitive response elicitor protein of the invention
XX
SQ Sequence 447 AA;

Query Match      100.0%; Score 2310; DB 5; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTIEQAQLLAELLSLS 60
DB 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTIEQAQLLAELLSLS 60
QY 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGCTGNNNSASSGTSSSGGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGCTGNNNSASSGTSSSGGSPFNDLSGGKA 180
QY 181 PSGNSPSGNSPVSTFTSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVGSAGIGAG 240
DB 181 PSGNSPSGNSPVSTFTSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQITFTAGSELGQSQENQKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQITFTAGSELGQSQENQKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGHLHYGDADKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
DB 301 LKNVTMGDDGADGHLHYGDADKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVTRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFVTRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420

RESULT 5
AAE18296
ID AAE18296 standard; protein; 447 AA.
XX
AC AAE18296;
XX
DT 07-MAY-2002 (first entry)
DE Erwinia amylovora hypersensitive response elicitor (HRE) #2.
XX
```

```
KW Hypersensitive response elicitor; HRE; transgenic plant; plant growth;
KW stress tolerance; disease tolerance; modified flower colour;
KW insect resistance; herbicide resistance; male sterility.
XX
OS Erwinia amylovora.
FN WO200195724-A2.
XX
PD 20-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US018955.
XX
PR 15-JUN-2000; 2000US-0211585P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Derocher J;
XX
WP1; 2002-130707/17.
DR N-PSDB; AAD29125.
XX
PT Improving effectiveness of transgenic plants by topical application of a
PT hypersensitive response elicitor protein to the transgenic plant or by
PT incorporating into the plant a transgene encoding the protein.
XX
PS Disclosure; Page 13-14; 86pp; English.
XX
CC The invention relates to methods of improving the effectiveness of
CC transgenic plants which involve either topical application of a
CC hypersensitive response elicitor (HRE) protein to the transgenic plant or
CC incorporating into the transgenic plant a transgene encoding HRE. HRE
CC sequence is used for improving the effectiveness of transgenic plants by
CC maximising the benefit of transgenic traits associated with a deleterious
CC effect on growth, stress tolerance, disease or insect resistance,
CC enhanced growth, herbicide resistance, male sterility, modified flower
CC colour and biochemically modified plant product in the transgenic plants
CC or overcoming the deleterious effects. The present sequence is Erwinia
CC amylovora HRE protein
XX
SQ Sequence 447 AA;

Query Match      100.0%; Score 2310; DB 5; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTIEQAQLLAELLSLS 60
DB 1 MSILTLNNTSSPGLFQSGDNGLGCHNANSALGQPIDRQTIEQAQLLAELLSLS 60
QY 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGVGNAGLNGRKGTAGTTPQSDSQNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGCTGNNNSASSGTSSSGGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGCTGNNNSASSGTSSSGGSPFNDLSGGKA 180
QY 181 PSGNSPSGNSPVSTFTSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVGSAGIGAG 240
DB 181 PSGNSPSGNSPVSTFTSPSTPTSPDPPSPPTKAAGGSTPTVTHDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQITFTAGSELGQSQENQKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGQITFTAGSELGQSQENQKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGHLHYGDADKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
DB 301 LKNVTMGDDGADGHLHYGDADKIDNLHVTNVGDEAITVKPNSAGKSHVEITNSSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVTRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFVTRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
```


Db	201	NPVRTSG-----TGSVSDTAGS-----TEOAGKINVVKTDTIKVGAGEVPDGHGAT	245
Qy	274	FTAGSELGGGSENOKPLFILEDGASLKQNTMGDDGADGIHLHG-----DAKIDNLHVTN	329
Db	246	FTADKSMGNGDQGENQKPMFELAEATLKQVNLGENEVDGIHVAKKAQEVTTIDNVHAQN	305
Qy	330	VGSDAITYKPSAGKSKSHVEIINSSFEHASDKILQLNADTNLSVDNVKAKDFGTFTVRTNG	389
Db	306	VGSDLTIVKGEGAAVTNLNIKNSSAKGADDKVQVLNANTHLKIDNFKADDFGTFTVRTNG	365
Qy	390	GQQ-GNWDNLNLSHISAEDGKPSFKVSDSEGLNVNTSDISLGDVENHY	435
Db	366	GKQFDMSIELNGIEANHGKFALVKSSDDLLKLTATGNTAMTDVKKAY	412
RESULT 9			
XX	AA71098		
ID	AA71098	standard; protein; 424 AA.	
XX	AA71098;		
XX	AC	XX	
XX	AC	XX	
DT	08-SEP-2000	(first entry)	
XX			
XX	Pseudomonas syringae dspE gene encoded hypersensitive response elicitor.		
DE			
XX			
XX	Hypersensitive response elicitor; environmental stress resistance; plant;		
KW	pathogen; dspE gene.		
KW			
XX			
OS	Pseudomonas syringae.		
XX			
PN	WO200028055-A2.		
XX			
PD	18-MAY-2000.		
XX			
PF	04-NOV-1999;	99WO-US026039.	
XX			
PR	05-NOV-1998;	98US-0107243P.	
XX			
PA	(EDEN-) EDEN BIOSCIENCE CORP.		
XX			
PI	Wei Z, Schading RL;		
XX			
DR	WPI; 2000-376566/32.		
DR	N-PSDB; AAD00673.		
XX			
PT	Application of a hypersensitive response elicitor protein to plants to		
PT	impart stress resistance.		
XX			
PS	Disclosure; Page 25-26; 84pp; English.		
XX			
CC	The patent discloses a method to impart stress resistance to plants by		
CC	applying a hypersensitive response elicitor in a non-infectious form to a		
CC	plant or seed. The present sequence is a hypersensitive response elicitor		
CC	protein encoded by dspE gene from Pseudomonas syringae. The protein		
CC	elicits a plant pathogen's hypersensitive response and is used to impart		
CC	stress resistance to plants		
XX			
SQ	Sequence 424 AA;		
Query Match 24.2%; Score 559; DB 3; Length 424;			
Best Local Similarity 36.1%; Pred. No. 2.7e-30;			
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;			
Qy	37	QPIDRQTIEQMALLAEALKSL---LSPQSGNAATGAGGNDQTTGVGNAGLNGRGTAG	93
Db	72	KPNDSQS--NIAKLISALIMSLQLMTNSNKKQDTNQEQPSQAPFQNGGLG-----	122
Qy	94	TTPQSQSONMLSEMGNNGLDQAITPDGCGGGGIGDNPILLKAMLKLIARMMDGSDQFGQP	153
Db	123	-TPSANS-----GGGG-----TPDATGGGG-GDTP-----SATGGG	151
Qy	154	GTGNASASGTSSSGSPFNDSLGGKAPSGNSPGNYSVPSTFTSPPTPTSPSPDLPFS	213

152	GGDTPTATGGGGGGGGTPTATGGG-----SGGTPTATGGGGGGVTPOITPQL-----A	200
214	SPTKAAGGSTPTVTDHPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITTVKAGQVFDGKGQT	273
201	NPNTSG-----TGVSVDTAGS-----TEQAGKINVVKDTIKVGAGEVFDGHGAT	245
274	PTAGSELGDGQSENQKPLFILEDGASLKQVMTMGDDGADGHIHLG-----DAKIDNLHVTN	329
246	PTADKSMGNGDQGENQKPMFELAEGATLKQVNLNGENEVDGIHVKAQNAQEVITIDNVHAQN	305
330	VGEDAITVKPNSAGKGSHEVLTNNSFEHASDKILQLNADTNLSVDNVKAKDFTCFVVRTNG	389
306	VGEDLIITVKGEGGAATVNLNTKNSAKAGDDKVVQQLNANTHUKIDNFKADDFGTWVRTNG	365
390	GQQ-GNMDNLNLSHISAEDGKFSFKVKSDEGLNVNTSDISLGDVENHY	435
366	GKQFDDMSIELNGIEANHKGKALVKSDSDLLKLATGNIAWTDVKHAY	412
RESULT 10		
AAV84859		
ID	AAV84859 standard; protein; 424 AA.	
XX		
AC	AAV84859;	
XX		
DT	08-AUG-2000 (first entry)	
XX		
DE	A hypersensitive response elicitor protein.	
XX		
KW	Hypersensitive response; insect control; disease resistance;	
KW	hypersensitive response elicitor; plant growth; vegetable; crop;	
KW	ornamental plant.	
XX		
OS	Pseudomonas syringae.	
XX		
PN	WO200020452-A2.	
XX		
PD	13-APR-2000.	
XX		
XX	05-OCT-1999; 99WO-US023181.	
PF		
XX		
PR	05-OCT-1998; 98US-0103050P.	
XX		
FA	(EDEN-) EDEN BIOSCIENCE CORP.	
XX		
PI	Wei Z, Pan H, Niggemeyer JL;	
XX		
DR	WPI; 2000-303745/26.	
DR	N-PSDB; AAA14943.	
XX		
PT	Hypersensitive response elicitor polypeptides useful for imparting	
PT	enhanced growth, disease resistance and insect resistance to plants,	
PT	especially vegetables and ornamental flowers.	
XX		
XX	Disclosure; Page 26-28; 100pp; English.	
XX		
CC	The present sequence represents a hypersensitive response elicitor	
CC	polypeptide. The specification describes hypersensitive response elicitor	
CC	polypeptide fragments, which do not elicit a hypersensitive response.	
CC	Instead, the proteins impart disease resistance to plants, enhance plant	
CC	growth, and/or control insects. The polypeptide fragments may be used to	
CC	these properties to plants. The plants which may be treated in this way	
CC	include vegetables, crops and ornamental plants such as alfalfa, rice,	
CC	wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet	
CC	potato, bean, pea, chichory, lettuce, endive, cabbage, brussel sprout,	
CC	beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,	
CC	garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,	
CC	cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,	
CC	pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis	
CC	thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,	
CC	carnation or zinnia	
XX		
XX	Sequence 424 AA;	
XX	50	

Sequence 424 AA;

XX PI Wei Z, Derocher J;
XX WPI; 2002-130707/17.
DR N-PSDB; AAD29127.
XX
XX Improving effectiveness of transgenic plants by topical application of a
PT hypersensitive response elicitor protein to the transgenic plant or by
PT incorporating into the plant a transgene encoding the protein.
XX
XX Disclosure; Page 17-18; 86pp; English.
XX
XX The invention relates to methods of improving the effectiveness of
CC transgenic plants which involves either topical application of a
CC hypersensitive response elicitor (HRE) protein to the transgenic plant or
CC incorporating into the transgenic plant a transgene encoding HRE. HRE
CC sequence is used for improving the effectiveness of transgenic plants by
CC maximising the benefit of transgenic traits associated with a deleterious
CC effect on growth, stress tolerance, disease or insect resistance,
CC enhanced growth, herbicide resistance, male sterility, modified flower
CC colour and biochemically modified plant product in the transgenic plants
CC or overcoming the deleterious effects. The present sequence is
CC Pseudomonas syringae HRE protein
XX
XX Sequence 424 AA;
XX
Query Match 24.2%; Score 559; DB 5; Length 424;
Best Local Similarity 36.1%; Pred. No. 2.7e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;
QY 37 QPIDRTIEMQAQLLAELKSL---LSQSGNAATGAGGNDQTTGVNAGGLNGRKGATG 93
DB 72 KPNDSQS--NIAKLISALIMSLQLMTNSNKKQDTNQEQPDSQAPFQNNGLG----- 122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPLLKAMKLARIWMDGSDQFGQP 153
DB 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNSASGTSSTSSGGSPFNDSLGGKAPSGNSPGNYSPVSTPTPTPTPTPTPTPTPTPT 213
DB 152 GGDTPATGGGGGGGGTPTATGGG---SGGTPATGGGGGVTPTQITPQL-----A 200
QY 214 SPTKAAGSTPTVTHDPDVGSAIGAGNSVAFVTSAGANQTVLHDTTVKAGQVFDKGQT 273
DB 201 NPNRTSG-----TGVSVDTAGS-----TEQAGKINVKDITKVGAGEVFDHGAT 245
QY 274 FTAGSELGGQSENOKPLFILEDGLASLKNVTMGDDGADGHIHYG----DAKIDNLHVTN 329
DB 246 FTADKSMGNGDQGENQKPFELAEAGATLKNVNGENEVDGIHVAKNAQEVITDINVHAQN 305
QY 330 VGEDAITVKNPNSAGKSHVEITNSSPFEHASDKILQNLADTNLSVDNVKAKDFGTFVRTNG 389
DB 306 VGEDLITVKEGGAATVNLNKNSSAKGADDKVQLNANTHLKIDNFKADDFGTFVRTNG 365
QY 390 GQO-GNWDNLNLSHSAEDKFSVKSDSEGLNVNTSDISLGDVENHY 435
DB 366 GKQFDDMSIELNGIEANHGKFALVKSDSDDLKATGNIAMTDVKHAY 412
RESULT 13
AAE16452
ID AAE16452 standard; protein; 424 AA.
XX
XX AAE16452;
XX
XX 09-APR-2002 (first entry)
DT
DE P. syringae hypersensitive response elicitor protein, depE.
XX
XX Hypersensitive response elicitor protein; plant growth; fruit coloration;
KW disease resistance; stress resistance; phytoalexin; insect infection;
XX plant maturation; depE protein.
XX

OS Pseudomonas syringae.
XX
XX Key Location/Qualifiers
FH Domain 45..102
FT /label= Hypersensitive_response_eliciting_domain
FT Region 45..79
FT /label= Acidic_unit
FT Region 79..102
FT /label= Alpha_helix
XX
XX WO200198501-A2.
XX
XX 27-DEC-2001.
XX
XX 12-JUN-2001; 2001WO-US018820.
XX
XX 16-JUN-2000; 2000US-0212211P.
XX
XX (EDEN-) EDEN BIOSCIENCE CORP.
XX
XX Fan H, Wei Z;
PI WPI; 2002-122282/16.
DR N-PSDB; AAD27020.
XX
XX New hypersensitive response elicitor proteins comprising spaced apart
PT domains having an acidic portion linked to an alpha-helix, useful for
PT imparting disease or stress resistance, controlling insects or enhancing
PT plant growth.
XX
XX Disclosure; Page 28-29; 99pp; English.
XX
XX The patent discloses hypersensitive response elicitor proteins and
CC nucleotides encoding such proteins. Hypersensitive response elicitor
CC proteins comprise an isolated pair or more of spaced apart domains, each
CC comprising an acidic portion linked to an alpha-helix and capable of
CC eliciting a hypersensitive response in plants. Sequences of the invention
CC are used to impart disease resistance to plants, to enhance plant growth,
CC to control insects and/or to impart stress resistance to plants which
CC includes resistance to environmental stresses such as climate, air
CC pollution, chemical and nutritional stress. The method of imparting
CC disease resistance has the potential for treating previously untreatable
CC diseases, treating diseases systemically and avoiding the use of
CC infectious agents or environmentally harmful materials. Hyper- sensitive
CC response elicitor sequences are used to enhance plant growth which
CC encompasses greater yield, increased in quantity of seeds produced,
CC percentage of seeds germinated, plant size and biomass, bigger fruits,
CC earlier fruit coloration and plant maturation. They are also used for
CC insect control which encompasses preventing direct insect damage to plant
CC by feeding injury, interfering with insect larval feeding on the plants,
CC preventing insects from colonising host plants and releasing phytotoxins.
CC Sequences of the invention also prevent subsequent disease damage to
CC plants resulting from insect infection. The present sequence is
CC Pseudomonas syringae hypersensitive response elicitor protein, depE
XX
XX Sequence 424 AA;
XX
Query Match 24.2%; Score 559; DB 5; Length 424;
Best Local Similarity 36.1%; Pred. No. 2.7e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;
QY 37 QPIDRTIEMQAQLLAELKSL---LSQSGNAATGAGGNDQTTGVNAGGLNGRKGATG 93
DB 72 KPNDSQS--NIAKLISALIMSLQLMTNSNKKQDTNQEQPDSQAPFQNNGLG----- 122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPLLKAMKLARIWMDGSDQFGQP 153
DB 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNSASGTSSTSSGGSPFNDSLGGKAPSGNSPGNYSPVSTPTPTPTPTPTPTPTPTPT 213
DB 152 GGDTPATGGGGGGGGTPTATGGG---SGGTPATGGGGGVTPTQITPQL-----A 200

CC protopectinase having an optimum pH 7 or higher against protopectin and
CC polygalacturonic acid substrates. Inclusion of protopectinase gives a
CC composition that provides better removal of muddy soil, particularly from
CC socks. The present sequence represents pectic acid lyase from *Bacillus*
CC sp. strain KSM-P15, which is used in an example from the present
CC invention. Pectic acid lyase exhibits protopectinase activity
XX
SQ Sequence 197 AA;

Query Match 15.6%; Score 359.5; DB 2; Length 197;
Best Local Similarity 46.3%; Pred. No. 6.2e-17;
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;
QY 253 TVLHDTITVKGQVFDGKGTFTAG-SELGDGQSQENQKPLFILEDGASLKNVTMGDDGA 311
Db 3 TVVHETIRVPAGQTFDGKGTFTAG-SELGDGQSQENQKPLFILEDGASLKNVTMGDDGA 62
QY 312 DGIHLYGDAKIDNLHVTNVEDAITVKPNSAGKSHVEITNSSEFHASDKILQADNTNL 371
Db 63 DGVHCVGDCITINWEDVGEDALTUK--SSGT---VNISGGAAYKAYDKVQFQINAAGTI 117
QY 372 SVDNVKAKDFGTFTVTRNGGQGNWDLNLSHISAEQKGFSEVKSQDSE---GLNVNT 423
Db 118 NIRNFRADDIGKLVQRNGGTTTKVVMNVENCNISRVKDAILRTDSTSTGRIVNT 172

RESULT 16
AAW77412
ID AAW77412 standard; protein; 197 AA.

XX AAW77412;
AC AAW77412;
XX 07-JAN-1999 (first entry)
DT 07-JAN-1999 (first entry)
XX
DE *Bacillus* sp. pectic acid lyase.
XX
KW *Bacillus* sp. KSM-P15; pectic acid lyase; pectinic acid lyase; pectin;
KW detergent; food-processing; fibre-processing agent; cell wall.
XX
OS *Bacillus* sp.

PN EP870834-A1.
XX 14-OCT-1998.
PD 14-OCT-1998.
XX 09-APR-1998; 98EP-00106586.
PF 09-APR-1998; 98EP-00106586.
XX 09-APR-1997; 97JP-00091142.
PR 09-APR-1997; 97JP-00091142.
XX 08-SEP-1997; 97JP-00242735.
PR 08-SEP-1997; 97JP-00242735.
XX (KAOS) KAO CORP.

XX Hatada Y, Koike K, Yoshimatsu T, Suzumatsu A, Kobayashi T, Ito S;
PI Hatada Y, Koike K, Yoshimatsu T, Suzumatsu A, Kobayashi T, Ito S;
XX WPI; 1998-523159/45.
DR N-PSDB; AAV59478.

XX New *Bacillus* pectic acid lyase - useful as a detergent component, a food-
PT processing agent and a fibre-processing agent.

XX Claim 1; Page 16-17; 29pp; English.

XX The present sequence represents a pectic acid lyase isolated from
CC microorganism *Bacillus* sp. KSM-P15. The pectic acid lyase has high pectic
CC acid lyase activity which degrades pectin in plant cell walls and fibre
CC in vegetables, and so is useful as a component of detergents, a food-
CC processing agent, or a fibre-processing agent. The pectic acid lyase has
CC a higher optimum reaction pH (10.3-10.7) than known *Bacillus* pectic acid
CC lyases (pH 8-9.5) and so has wider industrial applications. Unlike
CC present pectic acid lyases, the new enzyme has a high enzyme activity,
CC and can be produced on a mass scale

XX Sequence 197 AA;

Query Match 15.6%; Score 359.5; DB 2; Length 197;
Best Local Similarity 46.3%; Pred. No. 6.2e-17;
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;

QY 253 TVLHDTITVKGQVFDGKGTFTAG-SELGDGQSQENQKPLFILEDGASLKNVTMGDDGA 311
Db 3 TVVHETIRVPAGQTFDGKGTFTAG-SELGDGQSQENQKPLFILEDGASLKNVTMGDDGA 62
QY 312 DGIHLYGDAKIDNLHVTNVEDAITVKPNSAGKSHVEITNSSEFHASDKILQADNTNL 371
Db 63 DGVHCVGDCITINWEDVGEDALTUK--SSGT---VNISGGAAYKAYDKVQFQINAAGTI 117
QY 372 SVDNVKAKDFGTFTVTRNGGQGNWDLNLSHISAEQKGFSEVKSQDSE---GLNVNT 423
Db 118 NIRNFRADDIGKLVQRNGGTTTKVVMNVENCNISRVKDAILRTDSTSTGRIVNT 172

RESULT 17
AAW28446
ID AAW28446 standard; protein; 221 AA.
XX AAW28446;
AC AAW28446;
XX 17-JAN-2000 (first entry)
DT 17-JAN-2000 (first entry)
XX
DE *Bacillus* licheniformis Pectate lyase I.

XX Pectate lyase I; EC 4.2.2.2; pectin degrading enzyme; ATCC 14580;
KW catalytic active domain; cellulose binding domain; CBD; operably linked;
KW optimum activity; pH; detergent composition; yarn; cellulosic fibre;
KW recycled waste paper; pulp; retting process; animal feed; wine; juice;
KW transgenic plant.

XX *Bacillus* licheniformis.

XX Key Location/Qualifiers
FH Peptide 1..27
FT /label= Pro-sequence
FT Protein 28..221
FT /label= Mature_pectate_lyase_I
FT Misc-difference 133
FT /note= "Conserved residue"
FT Misc-difference 155
FT /note= "Conserved residue"

XX WO927083-A1.

XX 03-JUN-1999.

XX 24-NOV-1998; 98WO-DK000514.

XX 24-NOV-1997; 97DK-00001344.

PR 06-MAY-1998; 98US-00073684.

XX (NOVO) NOVO-NORDISK AS.

XX Andersen LN, Schuelein M, Lange NEK, Bjornvad ME, Schnorr K;

XX WPI; 1999-610578/52.

DR N-PSDB; AAX89484.

XX New isolated pectate lyase enzymes.

XX Claim 6; Page 79-80; 93pp; English.

XX The present sequence is pectate lyase I, which is a pectin degrading
CC enzyme derived from *Bacillus* licheniformis, ATCC 14580. The enzyme
CC comprises a catalytic active domain and a cellulose binding domain (CBD)
CC that are operably linked to each other. It shows optimum activity at pH
CC greater than 9 and temperature 55 degree centigrade. It can be used in
CC detergent compositions, for cleaning hard surfaces, for machine treatment
CC of fabrics, for improving the properties of cellulosic fibres, yarn,

CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1306 AA;

Query Match 8.6%; Score 199.5; DB 6; Length 1306;
 Best Local Similarity 26.0%; Pred. No. 8e-05;
 Matches 88; Conservative 16; Mismatches 122; Indels 113; Gaps 11;

QY 8 NNTSPGFLFQSGDNLGLGHNNALSALGQQPDRQTIEQMAQLLAKLLSPQSGNAA. 67
 DB 754 NGVAGSQPGGAGGCGTGGVCGNGRGIDGADGAT----- 789
 QY 68 TGAGGNDQTTGCVNAGLNGKCTAGTTTQSDQNMLSEMNGNLDQAITPDGQGGQIG 127
 DB 790 --AGARGQDGGAGGAGGKGRGTGP-----GGAGPAGTTGSGAGNG 832
 QY 128 DNPLLKAMLLIARMWDGSDQFGPCTGNNSASGTSSTSGSPFNDLSGKA-PSGNSP 186
 DB 833 -----GSGGTGDPDGGNGANGSVFTNNGTGGNGGNGAGPAGSGAGG 875
 QY 187 SGNYSPVSTPSTPTSP-----DFPSSPTKAAG-----G 221
 DB 876 SGAG--STFG---ATGSSSIHVNGGNGGNGGNGGNGGNGGNGGNGGNGGNGG 929
 QY 222 STPTVTHPPVDSAGTACAGNSVFTAGANQTVLHTITVKAQVFDGKGQTTAGSELG 281
 DB 930 SGAGGHHGNGGNASRGMGDDGTGTGAGGN-----AGQING-----GAGGNGG 973
 QY 282 DGQSGENKPLFTL-----EDGSLKNVTMGDDGDADG 313
 DB 974 DGTGSDGNPGLTSGSGRGDGGVGGGSGSVAGDGDG 1012

RESULT 20
 ABU36802
 ID ABU36802 standard; protein; 694 AA.
 AC ABU36802;
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by prokaryotic essential gene #22329.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO20027183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA40672.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 64726; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 694 AA;

Query Match 8.4%; Score 193; DB 6; Length 694;
 Best Local Similarity 25.4%; Pred. No. 0.0001;
 Matches 86; Conservative 28; Mismatches 140; Indels 84; Gaps 14;

QY 18 QSGDNLGLGH-NANSALGQQPDRQTIEQMAQLLAKLLSPQSGNNAATGAGNDQT 76
 DB 256 QTGGDGTGGHGTAGTPTGTGGDGT-----ATAGSKATGAGDGGT 300
 QY 77 TGVNAGLNGKRTAGTTTQSDQNMLSEMNGNLDQAITPDGQGGQIGDNPLLKAML 136
 DB 301 AAAGGCGGNGGAGVA-----QGDIAAFGCGDNGSDGVAAGSGGSGGAGG 352
 QY 137 KLIARMWDGSDQFGPCTGNNSASGTSSTSGSPFNDLSGKA-APSGNSPSGNYSPV 193
 DB 353 HIATATSTGSGGFG--GNGAASAGSAGDGGAGGAGGAGGLLPDGGNGGAGGAGGI 410
 QY 194 STFPSPTPTSPSTPLDPPSPSTKAAGSTPVT--DHDPD-----VGSAG 236
 DB 411 GG-----DGATGGPGSGGNAGIARFSDPPEAPDVVVGKGGGKGGSG 456
 QY 237 IGAGNSVATPSAGANQTVLHTITVKAQVFDGKGQTTAGSELGCGGSGENKPLFILE 296
 DB 457 LGVGCAGGTGCGAGNG-----GAGGLLFGNGN--GGNAGAGG-----D 493
 QY 297 DGASLKNVTMGDDGDGIHLYGDAKIDNL-HVTNVCED 333
 DB 494 GGAGVAGVGNGGCGGGGTATFHDPPVAGVWAVGGVGGD 531

Thu Jan 27 12:40:01 2005

us-09-879-248-6.rag

Page 14

Search completed: January 25, 2005, 12:06:04
Job time : 122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: January 25, 2005, 12:04:41 ; Search time 58 Seconds

(without alignments)
2784.420 Million cell updates/sec

Title: US-09-879-248-6

Perfect score: 2310

Sequence: 1 MSILLNNTSSPGLFQSG.....LGDVENVKVPMSANLKVAE 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	9	US-09-835-684-5
2	2310	100.0	447	9	US-09-880-371-5
3	2310	100.0	447	9	US-09-879-248-6
4	2310	100.0	447	14	US-10-010-390-5
5	2310	100.0	447	15	US-10-441-736-6
6	2310	100.0	447	17	US-10-847-142-5
7	559	24.2	424	9	US-09-835-684-9
8	559	24.2	424	9	US-09-880-371-9
9	559	24.2	424	9	US-09-879-248-14
10	559	24.2	424	14	US-10-010-390-9
11	559	24.2	424	15	US-10-441-736-14
12	559	24.2	424	17	US-10-847-142-9
13	256	11.1	276	14	US-10-156-761-13910

14	199.5	8.6	1306	15	US-10-282-122A-64405	Sequence 64405, A
15	193	8.4	594	15	US-10-282-122A-64726	Sequence 64726, A
16	191.5	8.3	584	15	US-10-282-122A-50604	Sequence 50604, A
17	190.5	8.2	226	17	US-10-425-115-229387	Sequence 229387, A
18	188	8.1	639	15	US-10-282-122A-64609	Sequence 64609, A
19	186.5	8.1	591	15	US-10-282-122A-64363	Sequence 64363, A
20	185.5	8.0	1381	15	US-10-282-122A-64895	Sequence 64895, A
21	185	8.0	355	17	US-10-425-115-313677	Sequence 313677, A
22	183	7.9	588	15	US-10-282-122A-64869	Sequence 64869, A
23	182	7.9	606	15	US-10-282-122A-64464	Sequence 64464, A
24	181.5	7.9	484	15	US-09-820-843A-19	Sequence 19, Appl
25	181.5	7.9	484	15	US-10-282-122A-64867	Sequence 64867, A
26	180.5	7.8	1079	10	US-09-820-843A-20	Sequence 20, Appl
27	176.5	7.6	562	15	US-10-282-122A-64514	Sequence 64514, A
28	174	7.5	615	15	US-10-282-122A-64786	Sequence 64786, A
29	174	7.5	2042	14	US-10-192-584-6	Sequence 6, Appl
30	172.5	7.5	584	15	US-10-282-122A-64903	Sequence 64903, A
31	172	7.4	3507	14	US-10-369-493-5784	Sequence 5784, Ap
32	171.5	7.4	667	15	US-10-282-122A-64494	Sequence 64494, A
33	171	7.4	505	15	US-10-282-122A-62341	Sequence 62341, A
34	169.5	7.3	914	15	US-10-282-122A-64606	Sequence 64606, A
35	169	7.3	525	15	US-10-282-122A-64763	Sequence 64763, A
36	167.5	7.3	1011	15	US-10-282-122A-64589	Sequence 64589, A
37	166.5	7.2	518	15	US-10-282-122A-50634	Sequence 50634, A
38	166.5	7.2	532	15	US-10-282-122A-64658	Sequence 64658, A
39	166	7.2	461	15	US-10-282-122A-64750	Sequence 64750, A
40	166	7.2	603	15	US-10-282-122A-64537	Sequence 64537, A
41	166	7.2	767	15	US-10-282-122A-48384	Sequence 48384, A
42	165.5	7.2	344	9	US-09-086-118-27	Sequence 27, Appl
43	165.5	7.2	344	9	US-09-835-684-11	Sequence 11, Appl
44	165.5	7.2	344	9	US-09-880-371-11	Sequence 11, Appl
45	165.5	7.2	344	9	US-09-879-248-15	Sequence 15, Appl
46	165.5	7.2	344	9	US-09-770-693-7	Sequence 7, Appl
47	165.5	7.2	344	9	US-09-766-348-7	Sequence 7, Appl
48	165.5	7.2	344	14	US-10-034-158-7	Sequence 7, Appl
49	165.5	7.2	344	14	US-10-010-390-11	Sequence 11, Appl
50	165.5	7.2	344	14	US-10-387-806-27	Sequence 27, Appl
51	165.5	7.2	344	14	US-10-441-736-15	Sequence 15, Appl
52	165.5	7.2	344	17	US-10-847-142-11	Sequence 11, Appl
53	163	7.1	618	15	US-10-282-122A-64608	Sequence 64608, A
54	162.5	7.0	1017	14	US-10-369-493-20166	Sequence 20166, A
55	161.5	7.0	576	15	US-10-282-122A-64547	Sequence 64547, A
56	161.5	7.0	837	15	US-10-282-122A-64362	Sequence 64362, A
57	161.5	7.0	1215	14	US-10-369-493-10073	Sequence 10073, A
58	161	7.0	1601	15	US-10-263-929-132	Sequence 132, App
59	161	7.0	1601	15	US-10-267-502-241	Sequence 241, App
60	160.5	6.9	957	15	US-10-282-122A-64361	Sequence 64361, A
61	159.5	6.9	778	15	US-10-282-122A-64751	Sequence 64751, A
62	159	6.9	714	14	US-10-032-585-7539	Sequence 7539, Ap
63	159	6.9	3286	15	US-10-282-122A-49697	Sequence 49697, A
64	157	6.8	491	15	US-10-282-122A-64558	Sequence 64558, A
65	157	6.8	979	15	US-10-282-122A-47372	Sequence 47372, A
66	156	6.8	2039	14	US-10-192-584-7	Sequence 7, Appl
67	155.5	6.7	2353	9	US-09-797-862-33	Sequence 33, Appl
68	155	6.7	1672	15	US-10-282-122A-55254	Sequence 55254, A
69	154.5	6.7	354	15	US-09-820-843A-21	Sequence 21, Appl
70	154.5	6.7	630	14	US-10-128-714-8597	Sequence 8597, Ap
71	154	6.7	334	15	US-10-282-122A-47939	Sequence 47939, A
72	154	6.7	1099	15	US-10-282-122A-48220	Sequence 48220, A
73	153.5	6.6	597	9	US-09-793-306-146	Sequence 146, App
74	153.5	6.6	597	15	US-10-320-797-3261	Sequence 3261, Ap
75	153	6.6	611	14	US-10-086-510-4	Sequence 4, Appl
76	153	6.6	1466	15	US-10-402-089-4	Sequence 4, Appl
77	153	6.6	1466	15	US-10-402-089-6	Sequence 6, Appl
78	153	6.6	1466	15	US-10-403-072A-4	Sequence 4, Appl
79	153	6.6	1466	15	US-10-402-072A-6	Sequence 6, Appl
80	153	6.6	1770	9	US-09-841-132-444	Sequence 444, App
81	153	6.6	1770	16	US-10-467-534-21	Sequence 21, Appl
82	153	6.6	1770	17	US-10-872-155-444	Sequence 444, App
83	152.5	6.6	580	16	US-10-367-094-3	Sequence 3, Appl
84	152.5	6.6	821	14	US-10-292-798-1370	Sequence 1370, Ap
85	152.5	6.6	852	15	US-10-282-122A-43144	Sequence 43144, A
86	152.5	6.6	923	15	US-10-282-122A-64474	Sequence 64474, A

87 152 6.6 469 15 US-10-282-122A-45275 Sequence 45275, A
88 152 6.6 1082 14 US-10-369-493-17890 Sequence 17890, A
89 152 6.6 3619 15 US-10-282-122A-67768 Sequence 67768, A
90 151 6.5 256 10 US-09-820-843A-18 Sequence 18, Appl
91 151 6.5 403 9 US-09-086-118-23 Sequence 23, Appl
92 151 6.5 403 9 US-09-835-684-3 Sequence 3, Appl
93 151 6.5 403 9 US-09-880-371-3 Sequence 3, Appl
94 151 6.5 403 9 US-09-879-248-3 Sequence 3, Appl
95 151 6.5 403 9 US-09-770-693-3 Sequence 3, Appl
96 151 6.5 403 9 US-09-766-348-3 Sequence 3, Appl
97 151 6.5 403 14 US-10-034-158-3 Sequence 3, Appl
98 151 6.5 403 14 US-10-010-390-3 Sequence 3, Appl
99 151 6.5 403 14 US-10-387-806-23 Sequence 23, Appl
100 151 6.5 403 15 US-10-441-736-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-835-684-5
; Sequence 5, Application US/09835684
; Patent No. US20020019337A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/71
; CURRENT APPLICATION NUMBER: US/09/835,684
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-835-684-5

Query Match 100.0%; Score 2310; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEMGNGLDQAITPDG 120
QY 121 QGGQIQDGNPLLKAMLKLIARMMDGQSDQFGQGTGNNSSASGTSSSGSGSPFNLDLGGKA 180
DB 121 QGGQIQDGNPLLKAMLKLIARMMDGQSDQFGQGTGNNSSASGTSSSGSGSPFNLDLGGKA 180
QY 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDFPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240
DB 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDFPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDQAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDQAS 300
QY 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360
DB 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420

RESULT 3

US-09-879-248-6
; Sequence 6, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min

Db 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
RESULT 2
US-09-880-371-5
; Sequence 5, Application US/09880371
; Patent No. US20020059558A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: DeRoche, Jay
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 21829/91
; CURRENT APPLICATION NUMBER: US/09/880,371
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/211,585
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-880-371-5

Query Match 100.0%; Score 2310; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIEQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGNDQTTGCVGNAGLNGRKTAGTTPOSQSONMLSEMGNGLDQAITPDG 120
QY 121 QGGQIQDGNPLLKAMLKLIARMMDGQSDQFGQGTGNNSSASGTSSSGSGSPFNLDLGGKA 180
DB 121 QGGQIQDGNPLLKAMLKLIARMMDGQSDQFGQGTGNNSSASGTSSSGSGSPFNLDLGGKA 180
QY 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDFPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240
DB 181 PSNGSPSGNYSPTVSTFSPSTPTSPPLDFPSSPTKAAGGSTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDQAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKGTFTAGSELGQSGQSENQKPLFILEDQAS 300
QY 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360
DB 301 LKNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHASD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFTVTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 3

US-09-879-248-6
; Sequence 6, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min

```

; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; FILE REFERENCE: THEREOF
; CURRENT APPLICATION NUMBER: US/09/879,248
; PRIOR FILING DATE: 2001-06-12
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-879-248-6

Query Match      100.0%; Score 2310; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIIDRQTIQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIIDRQTIQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
DB 61 POSGNAATGAGGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
QY 121 QGGGQIGDNPPLKAMUKLIARMMWGDQDQFGQGTGNNSSASCTSSSGGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPPLKAMUKLIARMMWGDQDQFGQGTGNNSSASCTSSSGGSPFNDLSGGKA 180
QY 181 PSGNSPSGNSPVSTSPSTPTSPPLDPSPTKAAAGSTPVTDHPDPVGSAGIGAG 240
DB 181 PSGNSPSGNSPVSTSPSTPTSPPLDPSPTKAAAGSTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKQGTFTAGSELGDCGQSENOKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKQGTFTAGSELGDCGQSENOKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 4
US-10-010-390-5
; Sequence 5, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; PRIOR FILING DATE: 2001-11-05
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora

US-09-879-248-6
```

```

US-10-010-390-5

Query Match      100.0%; Score 2310; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIIDRQTIQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIIDRQTIQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
DB 61 POSGNAATGAGGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
QY 121 QGGGQIGDNPPLKAMUKLIARMMWGDQDQFGQGTGNNSSASCTSSSGGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPPLKAMUKLIARMMWGDQDQFGQGTGNNSSASCTSSSGGSPFNDLSGGKA 180
QY 181 PSGNSPSGNSPVSTSPSTPTSPPLDPSPTKAAAGSTPVTDHPDPVGSAGIGAG 240
DB 181 PSGNSPSGNSPVSTSPSTPTSPPLDPSPTKAAAGSTPVTDHPDPVGSAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKQGTFTAGSELGDCGQSENOKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVDFGKQGTFTAGSELGDCGQSENOKPLFILEDGAS 300
QY 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
DB 301 LKNVTMGDDGADGIHLYGDAKIDNLHVTNVGDAITVKPNSACKKSHVEITNSSFEHASD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGTFVRTNGQQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; FILE REFERENCE: 21829/203 (BBC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6

Query Match      100.0%; Score 2310; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIIDRQTIQMAQLLAELLKSLLS 60
DB 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIIDRQTIQMAQLLAELLKSLLS 60
QY 61 POSGNAATGAGGNDQTTGVGNAGLNGRKGTAGTTTQSDSQNNMLSEMGNNGLDQAITPDG 120
```

Db 61 POSGNAATGAGGNDQTTGCVNAGGLNGRKGTTAGTTPOSDSQNMLSEMGNGGLDQAITPDG 120
Qy 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGTGNNSASGTSSSGGSPFNDLSGGKA 180
Db 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGTGNNSASGTSSSGGSPFNDLSGGKA 180
Qy 181 PSGNSPSGNYSPVSTFSPPTPTGTPPLDPPSSPTKAAAGGSTPVTDHPDPVGSAGIGAG 240
Db 181 PSGNSPSGNYSPVSTFSPPTPTGTPPLDPPSSPTKAAAGGSTPVTDHPDPVGSAGIGAG 240
Qy 241 NSVAFTSAGANQTVLHDTITVKAGOVFDGKGTTFAGSELGDCGQSENOKPLFILEDGAS 300
Db 241 NSVAFTSAGANQTVLHDTITVKAGOVFDGKGTTFAGSELGDCGQSENOKPLFILEDGAS 300
Qy 301 LKQVNTMGDDGADGIHLYGDAKIDNLHVTNVGDEDAITVKPNSAGKKSHEVITNSSFHASD 360
Db 301 LKQVNTMGDDGADGIHLYGDAKIDNLHVTNVGDEDAITVKPNSAGKKSHEVITNSSFHASD 360
Qy 361 KILQINADTNLSVDNVKAKDFGTFTVNTGGQGNWDLNLSHISAEDGKFSFKVSDSEGLN 420
Db 361 KILQINADTNLSVDNVKAKDFGTFTVNTGGQGNWDLNLSHISAEDGKFSFKVSDSEGLN 420
Qy 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 6
US-10-847-142-5
; Sequence 5, Application US/10847142
; Patent No. US20040265442A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/197
; CURRENT APPLICATION NUMBER: US/10/847,142
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/835,684
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-847-142-5

Query Match 100.0%; Score 2310; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 1.4e-151;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSILTLNNNTSSPGLFQSGGDNGLGCHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
Db 1 MSILTLNNNTSSPGLFQSGGDNGLGCHNANSALGQPIDRQTIQMAQLLAELLKSLLS 60
Qy 61 POSGNAATGAGGNDQTTGCVNAGGLNGRKGTTAGTTPOSDSQNMLSEMGNGGLDQAITPDG 120
Db 61 POSGNAATGAGGNDQTTGCVNAGGLNGRKGTTAGTTPOSDSQNMLSEMGNGGLDQAITPDG 120
Qy 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGTGNNSASGTSSSGGSPFNDLSGGKA 180
Db 121 QGGGQIGDNPLLKAMLLIARMMDGQSDQFQPGTGNNSASGTSSSGGSPFNDLSGGKA 180
Qy 181 PSGNSPSGNYSPVSTFSPPTPTGTPPLDPPSSPTKAAAGGSTPVTDHPDPVGSAGIGAG 240
Db 181 PSGNSPSGNYSPVSTFSPPTPTGTPPLDPPSSPTKAAAGGSTPVTDHPDPVGSAGIGAG 240

RESULT 7

US-09-835-684-9
; Sequence 9, Application US/09835684
; Patent No. US20020019337A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/71
; CURRENT APPLICATION NUMBER: US/09/835,684
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-835-684-9

Query Match 24.2%; Score 559; DB 9; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

Qy 37 QPIDRQTIQMAQLLAELLKSL---LSPQSGNAATGAGGNDQTTGCVNAGGLNGRKGTTAG 93
Db 72 KPDNSQS---NIAKLISALIMSLLOMLTNSNKKQDTNQOPDSQAPFQNNGLG----- 122
Qy 94 TTPQSDSQNMLSEMGNGGLDQAITPDGQGGGIGDNPLLKAMLLIARMMDGQSDQFGQP 153
Db 123 -TPSADS-----GGGG---TPDATGGG-GDTP-----SATGGG 151
Qy 154 GTGNNSASGTSSSGGSPFNDLSGGKAPSGNSPSGNYSPVSTFSPPTPTGTPPLDPPS 213
Db 152 GDTPTATGGGSGGGGTTATGGG---SGGTPTATGGGEGGVTTPQITPOL-----A 200
Qy 214 SPTKAAAGGSTPVTDHPDPVGSAGIAGNSVAPTSAGANQTVLHDTITVKAGOVFDGKGT 273
Db 201 NPNRTSG-----TGVSVDTAGS---TEAGKINVVKDTIKVGADEVDFGHGAT 245
Qy 274 FTAGSELGDCGQSENOKPLFILEDGASLKQVNTMGDDGADGIHLYG---DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENQKPMFELAEATLKNVNLGENEVDGIHVAKAKNAQEVTTIDNVHAQN 305
Qy 330 VGEDAITVKPNSAGKKSHEVITNSSFHASDKILQINADTNLSVDNVKAKDFGTFTVNTNG 389
Db 306 VGEDLITVKGEGGAATVNLINKNSAKGADKKVQVLNANTHLKIDNFKADDFGTMVTRNG 365
Qy 390 GQQ-GNMDNLNLSHISAEDGKFSFKVSDSEGLNVTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDSDDLKLAATGNIAMTVDVKHAY 412


```
RESULT 8
US-09-880-371-9
; Sequence 9, Application US/09880371
; Patent No. US20020059658A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Derocher, Jay
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 21829/91
; CURRENT APPLICATION NUMBER: US/09/880,371
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/211,585
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-880-371-9

Query Match 24.2%; Score 559; DB 9; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELLKSL---LSPQSGNAATGAGNDQTTGVGNAGGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQLMLTNSNKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLIARMMDGSDQDFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNASSTSSGGSPFNDLSGKAPSGNSGNYSPVSTFSPSTPTSPSPDLPFS 213
Db 152 GGDTPATGGGGGGGTPATGGG---SGTPTATGGGEGGVTPTTQPOL-----A 200
QY 214 SPTKAAAGSTPVTDPVPGSAGTACGNSVAFSTAGANQTVLHDTITVKAGQVDFGKGOT 273
Db 201 NPRTSG-----TGVSVDTAGS-----TEQAGKINNVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTWGGDAGDIHLYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENQKPMPELAEGATLKNVNLGENEVDGIHVAKAKNAQEVTTIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDSDDLKATGNIAMTDVKHAY 412

RESULT 9
US-09-879-248-14
; Sequence 14, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 21829/81
; CURRENT APPLICATION NUMBER: US/09/879,248
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/212,211
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14

Query Match 24.2%; Score 559; DB 14; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELLKSL---LSPQSGNAATGAGNDQTTGVGNAGGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQLMLTNSNKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLIARMMDGSDQDFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNASSTSSGGSPFNDLSGKAPSGNSGNYSPVSTFSPSTPTSPSPDLPFS 213
Db 152 GGDTPATGGGGGGGTPATGGG---SGTPTATGGGEGGVTPTTQPOL-----A 200
QY 214 SPTKAAAGSTPVTDPVPGSAGTACGNSVAFSTAGANQTVLHDTITVKAGQVDFGKGOT 273
Db 201 NPRTSG-----TGVSVDTAGS-----TEQAGKINNVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTWGGDAGDIHLYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENQKPMPELAEGATLKNVNLGENEVDGIHVAKAKNAQEVTTIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDSDDLKATGNIAMTDVKHAY 412
```

```
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-879-248-14

Query Match 24.2%; Score 559; DB 9; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELLKSL---LSPQSGNAATGAGNDQTTGVGNAGGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQLMLTNSNKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLIARMMDGSDQDFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNASSTSSGGSPFNDLSGKAPSGNSGNYSPVSTFSPSTPTSPSPDLPFS 213
Db 152 GGDTPATGGGGGGGTPATGGG---SGTPTATGGGEGGVTPTTQPOL-----A 200
QY 214 SPTKAAAGSTPVTDPVPGSAGTACGNSVAFSTAGANQTVLHDTITVKAGQVDFGKGOT 273
Db 201 NPRTSG-----TGVSVDTAGS-----TEQAGKINNVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTWGGDAGDIHLYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENQKPMPELAEGATLKNVNLGENEVDGIHVAKAKNAQEVTTIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDSDDLKATGNIAMTDVKHAY 412

RESULT 10
US-10-010-390-9
; Sequence 9, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/248,169
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-010-390-9

Query Match 24.2%; Score 559; DB 14; Length 424;
Best Local Similarity 36.1%; Pred. No. 1.9e-30;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELLKSL---LSPQSGNAATGAGNDQTTGVGNAGGLNGRKGATAG 93
Db 72 KPNDQS--NIAKLISALIMSLQLMLTNSNKKQDTNQEOPDSQAPFQNGGLG-----122
QY 94 TTPQSDSQNMLSEMGNGLDQAITPDGCGGQIGDNPPLKAMKLIARMMDGSDQDFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
QY 154 GTGNNASSTSSGGSPFNDLSGKAPSGNSGNYSPVSTFSPSTPTSPSPDLPFS 213
Db 152 GGDTPATGGGGGGGTPATGGG---SGTPTATGGGEGGVTPTTQPOL-----A 200
QY 214 SPTKAAAGSTPVTDPVPGSAGTACGNSVAFSTAGANQTVLHDTITVKAGQVDFGKGOT 273
Db 201 NPRTSG-----TGVSVDTAGS-----TEQAGKINNVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTWGGDAGDIHLYG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNGDQGENQKPMPELAEGATLKNVNLGENEVDGIHVAKAKNAQEVTTIDNVHAQN 305
QY 330 VGEDAITVKPNSAGKSHVEITNSSPEHASDKILOLNADTNLSVDNVKAKDFGTFVVRTNG 389
Db 306 VGEDLITVKGEGGAATVNLNKNSSAKGADDKVVQLNANTHLKIDNFKADDFGTFVVRTNG 365
QY 390 GQO-GNWDNLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLGDVENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFAVKSDSDDLKATGNIAMTDVKHAY 412
```



```
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64363
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64363

Query Match      8.1%; Score 186.5; DB 15; Length 591;
Best Local Similarity 23.1%; Pred. No. 0.00017;
Matches 91; Conservative 31; Mismatches 119; Indels 153; Gaps 15;

QY 6 LNNNTSSPGLFQSGDNGL-----GHNANSALGQPIDRQTIEQMAQLLAEKLSLSP 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 IGGANGAGPGTGQAGGDLGLFNGNGGSGAPQ-----AG 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 QSGNAA-----TGAGNDOTTGVGNAG----- 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 GAGGAAGFFNGNGGSGGAGAGGAGGTAGWFFGNGGAGGIGVAGINGLGGAGGD 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 -----GLNGRKGTAG-----TTPOSDSQNMLSENGNGLDQAITPDGQGG 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 GGNAGFFNGNGGSGGAGAGVAVNPGLATPVPAAN-----GGNGLLVGVPTAGG 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 GQIGDNPLKAMKLIIARMMDGSDQFGPQGTGNNSSAG-----TSSGSGSPFNDLSG 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 GADGAN-----GSAIGQAG--GAGGDDGNASTSGGIGIAQTGGAGGA--GGAGG 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 GKAPSGNSPSGNSYSPVSTFPSPPTSPSTPLDFPSPPTKAAAGSTPVTDHPDPVGSAGI 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 DGAPGGNGNG-----GSVEHTGATGSSAGSGNGATGGNGGVGAPG- 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 238 GAGNSVAFTSAGANQTVLHDTITVKAQGVFDGKQGTFTAGSELGDCGQSENKPLFILED 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 GAGGNGHVSGSVNT-----AGA--GGKGGNGTGGAGCPGCH----- 390
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 GASLKNVTMGDGDGADGHLHYGDADKIDNLHVTNNGV 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 391 GGSVLSGVPVDSGNGGAGGSGGAGVSATDIAGTG 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 20
US-10-282-122A-64895
; Sequence 64895, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
```

```
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64895
; LENGTH: 1381
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64895

Query Match      8.0%; Score 185.5; DB 15; Length 1381;
Best Local Similarity 25.0%; Pred. No. 0.00058;
Matches 79; Conservative 25; Mismatches 141; Indels 71; Gaps 11;

QY 19 SGGDNGLGCHNANSALGQQ--PIDRQTIEQMAQLLAEKLSLSPQSGNAATGAGNDQT 76
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 779 SGGDGGKGGGSGGTGGSGAPI-----GGGAGGTGGSGGH 814
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 TGVNAGLNGRKGTAGTTPQSDSQNMLSENGNGLDQAITPDGQGG-GQIGDNPLLKAM 135
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 815 AGKGGAGGI-GAQGTITIVPNGGN--AGDGGNGNAGA---CGNGSGDFFGNT----- 863
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 LKLIIARMMDGSDQFGPQGTGNNSSAGSTSSSGSGSPFNDLSGCKAPS-GNSPSGNSYSPVS 194
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 864 -----TSGASGSGGNGNAGTAGSGGAGGTGTGLSGGNGGNGGNGGNGGNGG 913
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 195 TFSPPSTPSTPLDFPSPPTKAAAGSTPVTDHPDPVGSAGIAGNSVAFTSAGANQTV 254
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 AHGTVGAQGFVPATSLPTFNGGAGGNGGTGSGNGGAPGAPAGPTTGGNAGSQGIGGGGNG 973
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 255 LHDITIVKAGVFDG-----KGQTTFTAGS-----ELGDCGQSENKPLFILED 297
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 974 GGD-----GGKGGGDADAVNVVFMPTBFOAATGTAGSAGDPTGNGGPGTGPSPMVAPPP 1028
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 GASLKNVTMGDDGADG 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1029 PTPITVQQGGDGGAG 1044
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: January 25, 2005, 12:09:46
Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:02:31 ; Search time 38 Seconds

(without alignments)
1131.813 Million cell updates/sec

Title: US-09-879-248-6

Perfect score: 2310

Sequence: 1 MSLLTNNNTSSPGLFQSG.....LGDVNHVKVPMANLKVAE 447

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2310	100.0	447	2 T18447	HrpW protein - Erw
2	353.5	15.3	221	2 A70045	pectate lyase homo
3	233.5	10.1	219	2 S68364	pectate lyase (EC
4	207	9.0	1660	2 A70869	hypothetical glyci
5	199.5	8.6	1306	2 A70934	hypothetical glyci
6	193	8.4	694	2 F70868	hypothetical glyci
7	188	8.1	639	2 D70931	hypothetical glyci
8	186.5	8.1	591	2 B70523	hypothetical glyci
9	186	8.1	1329	2 E70917	hypothetical glyci
10	185.5	8.0	1381	2 E70806	hypothetical glyci
11	185.5	8.0	1489	2 D70807	pectate lyase (EC
12	183.5	7.9	242	2 A45724	hypothetical glyci
13	183	7.9	588	2 F70971	hypothetical glyci
14	182	7.9	606	2 H70816	hypothetical glyci
15	182	7.9	731	2 C70874	hypothetical glyci
16	181.5	7.9	484	2 G70846	hypothetical glyci
17	180.5	7.8	1079	2 B70807	hypothetical glyci
18	179	7.7	730	2 F96559	hypothetical prote
19	179	7.7	1538	2 B70846	hypothetical glyci
20	179	7.7	3016	2 S77300	hypothetical prote
21	178	7.7	741	2 G70917	hypothetical glyci
22	177.5	7.7	1901	2 F70806	hypothetical glyci
23	177.5	7.7	2232	2 T34434	hypothetical prote
24	176.5	7.6	562	2 B70953	hypothetical glyci
25	175	7.6	645	2 F70825	probable PPE prote
26	174	7.5	615	2 H70589	hypothetical glyci
27	172.5	7.5	584	2 G70804	hypothetical glyci
28	172	7.4	3507	2 T34513	hypothetical prote
29	171.5	7.4	667	2 A70893	hypothetical glyci

30	171.5	7.4	1778	2 T50074	probable nucleopor
31	171	7.4	463	2 B70893	hypothetical glyci
32	169.5	7.3	514	2 H70987	hypothetical glyci
33	169	7.3	525	2 D70878	hypothetical glyci
34	168.5	7.3	398	1 OZQAS	circumsporozoite p
35	167.5	7.3	1011	2 F70620	hypothetical glyci
36	167.5	7.3	2174	2 E95965	hypothetical glyci
37	166.5	7.2	532	2 F70580	hypothetical glyci
38	166.5	7.2	714	2 A70807	hypothetical glyci
39	166	7.2	461	2 F70571	hypothetical glyci
40	166	7.2	603	2 A70770	hypothetical glyci
41	166	7.2	1207	2 T23754	hypothetical prote
42	165.5	7.2	344	2 S41707	PopAI protein - Ps
43	165	7.1	591	2 S74999	iron-regulated pro
44	165	7.1	749	2 A70812	hypothetical glyci
45	165	7.1	959	2 B44402	nuclear pore compl
46	164	7.1	767	2 B70895	hypothetical glyci
47	164	7.1	839	2 H90577	lipoprotein vsa1 f
48	163	7.1	618	2 A70989	hypothetical glyci
49	162	7.0	434	2 E70768	hypothetical glyci
50	161.5	7.0	576	2 A70900	hypothetical glyci
51	161.5	7.0	837	2 E70835	hypothetical glyci
52	161	7.0	787	2 T00798	hypothetical prote
53	160.5	6.9	314	2 F70766	hypothetical prote
54	160.5	6.9	853	2 A70896	hypothetical glyci
55	160.5	6.9	937	2 S58135	hypothally regulated
56	160.5	6.9	957	2 D70835	hypothetical glyci
57	159.5	6.9	778	2 F70963	hypothetical glyci
58	159.5	6.9	786	2 T16509	hypothetical prote
59	159	6.9	286	2 C61615	sericin MG-2 - gre
60	159	6.9	1341	2 H98123	hypothetical prote
61	158.5	6.9	1275	2 T33369	hypothetical prote
62	157.5	6.8	2329	2 T28125	hypothetical prote
63	157	6.8	491	2 D70916	hypothetical glyci
64	157	6.8	543	2 F70726	hypothetical glyci
65	156.5	6.8	496	2 H70839	hypothetical glyci
66	156.5	6.8	1428	2 T08852	lustrin A - Califo
67	155.5	6.7	572	2 T08509	trbl protein - Ent
68	155	6.7	1113	2 S28925	nuclear pore compl
69	155	6.7	1672	2 C81675	polymorphic membra
70	155	6.7	2249	2 A41477	190K surface antig
71	155	6.7	3198	2 A43426	collagen alpha 2 f
72	154.5	6.7	354	2 B70563	probable PPE prote
73	154.5	6.7	582	2 F70675	probable PPE prote
74	154	6.7	403	2 T08471	harpin - Erwinia a
75	154	6.7	430	2 S52700	NUP42 protein - ye
76	153.5	6.6	590	2 E70946	probable PPE prote
77	153	6.6	788	2 JS0747	regulatory protein
78	153	6.6	1049	1 CGB075	collagen alpha 1(I
79	153	6.6	1770	2 A71517	hypothetical prote
80	152.5	6.6	552	2 D70604	probable PPE prote
81	152.5	6.6	683	2 A82704	1,4-beta-cellobios
82	152.5	6.6	923	2 E70820	hypothetical glyci
83	152.5	6.6	2020	2 C48399	ABC-type transport
84	152.5	6.6	3190	2 T13828	CREB-binding prote
85	152	6.6	1560	2 T02885	peroxisome prolife
86	151	6.5	256	2 A70514	hypothetical glyci
87	151	6.5	395	2 A41156	circumsporozoite p
88	151	6.5	1190	2 A82615	surface protein XF
89	151	6.5	1844	2 T51890	related to Nup98-N
90	151	6.5	2554	2 AB3528	extracellular seri
91	150.5	6.5	598	2 T38403	probable nucleopor
92	150.5	6.5	1414	1 S23809	collagen alpha 2(I
93	150	6.5	515	2 H70663	hypothetical glyci
94	150	6.5	1070	2 A10484	probable autotrans
95	150	6.5	1407	2 B72078	polymorphic outer
96	149.5	6.5	2340	2 B71704	cell surface antig
97	149	6.5	605	2 T33913	hypothetical prote
98	149	6.5	681	2 AB2155	hypothetical prote
99	148.5	6.4	1028	2 T03516	probable outer mem
100	148	6.4	387	2 C41156	circumsporozoite p

ALIGNMENTS

RESULT 1

T18447

HrpW protein - Erwinia amylovora

C:Species: Erwinia amylovora

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18447

R:Gaudriault, S.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z18936

A:Accession: T18447

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-447 <GAU>

A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:CAA741

A:Experimental source: strain CFBP1430; specific host Pomoideae

C:Genetics:

A:Note: hrpW

Query Match 100.0%; Score 2310; DB 2; Length 447;
 Best Local Similarity 100.0%; Pred. No. 3.3e-119; Mismatches 0; Indels 0; Gaps 0;
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGIFOSGGDNGLGGHNANSALGQQPIDRQTIEQMAQLLAELKSLLS 60
 DB 1 MSILTLNNNTSSPGIFOSGGDNGLGGHNANSALGQQPIDRQTIEQMAQLLAELKSLLS 60

QY 61 POSGNAATGAGGNDQTTGVNAGGLNGKRTAGTTTQSDSONMLSEMGNGLDQAITPDG 120
 DB 61 POSGNAATGAGGNDQTTGVNAGGLNGKRTAGTTTQSDSONMLSEMGNGLDQAITPDG 120

QY 121 QGGGIGDNPPLKAMKLIARMDQSDQFQPGGTGNNSSSGTSSSGSPFNDLSGKA 180
 DB 121 QGGGIGDNPPLKAMKLIARMDQSDQFQPGGTGNNSSSGTSSSGSPFNDLSGKA 180

QY 181 PSNGSPSGNYSPVSTFSPSTPTSPLOPPSPPTKAAGSTPTVTHPPVPVSGAGIGAG 240
 DB 181 PSNGSPSGNYSPVSTFSPSTPTSPLOPPSPPTKAAGSTPTVTHPPVPVSGAGIGAG 240

QY 241 NSVAFPSAGANQTVLHDTITVKAQGVDFGKGQTFAGSELGQSQSNOKPLFILEDGAS 300
 DB 241 NSVAFPSAGANQTVLHDTITVKAQGVDFGKGQTFAGSELGQSQSNOKPLFILEDGAS 300

QY 301 LKNTMTGDDGADGTHLYGDAKIDNLHVTNVEDAITVKPNSAGKSHVEITNSFPHASD 360
 DB 301 LKNTMTGDDGADGTHLYGDAKIDNLHVTNVEDAITVKPNSAGKSHVEITNSFPHASD 360

QY 361 KILQNLADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLN 420
 DB 361 KILQNLADTNLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLN 420

QY 421 VNTSDISLGDVENHYKVPNSANLKVAE 447
 DB 421 VNTSDISLGDVENHYKVPNSANLKVAE 447

RESULT 2

A70045

pectate lyase homolog yypA - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: A70045

R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez

C.; Bron, S.; Brouillet, C.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: A70045
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-221 <KUN>
 A:Cross-references: UNIPROT:O34310; GB:Z99121; GB:AL009126; NID:92635827; PIDN:CAB15500 1;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yypA

Query Match 15.3%; Score 353.5; DB 2; Length 221;
 Best Local Similarity 41.8%; Pred. No. 1e-12; Mismatches 31; Indels 5; Gaps 1;
 Matches 76; Conservative 31; Mismatches 31; Indels 5; Gaps 1;

QY 250 ANQTVLHDTITVKAQGVDFGKGQTFAGSELGQSQSNOKPLFILEDGASLKNVTMGDD 309
 DB 27 AADKVHETIIVPKNTTYDGKQRFVAGKELGDSQSENQDPVPRVEDGATLKVVILGAP 86

QY 310 GADGTHLYGDAKIDNLHVTNVEDAITVKPNSAGKSHVEITNSFPHASDKILOINADT 369
 DB 87 AADGVHTYGNVNIQNVKEDVGEDALTVK-----KEGKVTIDGSAQKASDKIFQINKAS 141

QY 370 NLSVDNVKAKDFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLNVNTSDISLG 429
 DB 142 TPTVKNFATADNGKFRQLGGSTFHVVDIADKITTNNKEALFRDTSKTSVTMTNTYS 201

QY 430 DV 431
 DB 202 NV 203

RESULT 3

S68364

pectate lyase (EC 4.2.2.2) C - fungus (Fusarium solani)

C:Species: Fusarium solani

C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S68364

R:Guo, W.; Gonzalez-Candelas, L.; Kolattukudy, P.E.

Arch. Biochem. Biophys. 323, 352-360, 1995

A:Title: Cloning of a new pectate lyase gene pelC from Fusarium solani f. sp. pisi (Nectr

A:Reference number: S68364; MUID:96063610; PMID:7487098

A:Accession: S68364

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-219 <GUO>

A:Cross-references: UNIPROT:Q00843; EMBL:U13049

C:Genetics:

A:Gene: pelC

A:Introns: 52/1; 102/3

C:Keywords: carbon-oxygen lyase

Query Match 10.1%; Score 233.5; DB 2; Length 219;
 Best Local Similarity 30.1%; Pred. No. 3.6e-06;
 Matches 63; Conservative 39; Mismatches 72; Indels 35; Gaps 6;

QY 220 GGSTPTVTHDPPVSGAGIGAGNSVAFTSAGANQTVLHDTITVKAQGVDFGKGQTFAGSE 279
 DB 8 GGVPKPTDH-----ISNS-----KVIEKAGQVYDGKWKAYDRGSG 43

QY 280 LGDG-QGSENOKPLFILEDGASLKNVTMGDDGADGTHLYGDAKIDNLHVTNVEDAITVK 338
 DB 44 ACKQNEGGDKDAVFLIHGATLKVNIIKDSGEGVHCCHGCHTLEFPWFEDVCEDAISIK 103

QY 339 PNSAGKSHVEITNSFPHASDKILOINADTNLSVDNVKAKDFGTFVRT--NGGQGNWD 396
 DB 104 EDKAGKESW--ITGGAYHASDKVVGNGCGTWNINFFVEDYKLYRSCGNSCKQCKRN 161

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70974
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-731 <COL>
A/Cross-references: GB:AL123456; GB:AL123456; NID:g3242262; PIDN:CAA15773.1; PID:g266166
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: Rv3388
C/Superfamily: elastin related uncharacterized glycine-rich protein, PE motif containing

Query Match 7.9%; Score 182; DB 2; Length 731;
Best Local Similarity 25.3%; Pred. No. 0.01;
Matches 89; Conservative 26; Mismatches 127; Indels 110; Gaps 15;

Qy 7 NNNTSSPGLFQSOGDNLGCHNANSALGOQPIDRTIQEQAQLLELLKSLSPQSGNA 66
Db |||:::|||||
399 NNTSAGTGGVAGSGGTGNAG-----LIGAGHGG 428

Qy 67 ATGAGGNDQTTVGVNAGLGKRKTAGTTTPQ-----SDSQN-----MLSEMGNNGLDQAIT 117
Db |||:::|||||
429 AGGAGG-N-QTGVCN-GGAGNGGAGGAGGQLYNGGDGGNGGAGGANIAGNGSDGGA 486

Qy 118 PDGCGGQIGDNPLLKAMLLIARMDDGQSDQFGPGCTGNNSAS-----SGTSSGGSP 171
Db |||:::|||||
487 GHGAGGS-----ARLI-GAGHGCGGAGGNTAGRADAATAGTCGGGNG 531

Qy 172 FNDLSGCKAPSGNSPSNYSPVSTFTSPPTPTPTPLDFPSSPTKAAGGSTPVTHDPDP 231
Db |||:::|||||
532 GN-----GCLLSGNAGAGGHGAGGSATTITTG-----TPPTGATGN----- 570

Qy 232 VGSAGIGAGNSVAFTSAGANQTVLHDITVKAGQVFHGKGOTFTAGSELCDGCQSENQKP 291
Db |||:::|||||
571 -GGNG-GAGGTAGTGG-----GICGNGAGGTGNGAVA----- 604

Qy 292 LFILEDGASLNVTMBDGDADGHIHYGDAKIDNLHVNVGEDAITVKPNASG 343
Db |||:::|||||
605 LSVGSTGLGNGSGSLGCGGSLFCNGGAGVGATG-GNGGSGIGPASVG 655

RESULT 16
G70846
hypothetical glycine-rich protein Rv3344c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C/Accession: G70846
R/Cole, S.F.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: G70846
A/Status: preliminary;
A/Molecule type: DNA
A/Residues: 1-484 <COL>
A/Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAA17116.1; PID:g289425
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: Rv3344c
C/Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 7.9%; Score 181.5; DB 2; Length 484;
Best Local Similarity 26.2%; Pred. No. 0.0065;
Matches 90; Conservative 24; Mismatches 108; Indels 121; Gaps 19;

Qy 14 PGLFQSGDNLGCHNANSALGOQPIDRTIQEQAQLLELLKSLSPQSGNAAT---G 69
Db PG--STGGAGKGGAGANG-----SSNGNRAGNAGNGHGG 99

Qy 70 AGGNDQTTVGVNAGLGKRKTAGTTTPQSDSQNNLMSEMGNNGLDQAITPDGQG----- 122

```

Db      100 AGSGDTCGAGAGGCGGFGGTGSG-----GSGIGGAGGNG-----GNGGAGGTGV 146
Qy      123 --GQIGDNPILLKAMLKLIARMMDGQSDQFG---OPGTGNNASASSCTSSSGSPFNDL-S 176
Db      147 VLGGKGGDG-----GNGDHGGPATNPWGSRGAGCGSGGNGGAGGNATGS 191
Qy      177 GKG-APSGNSPGSNYPVSTPSPSTPTSPSPPLDPSPSP-----TKAAGGSTPVTVDHPD 230
Db      192 GKGAGGNGGDSFG-----ATSGPASIGVTGAPGNGGKGGAGGS-----N 234
Qy      231 PVGSAGTG--AGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGQPTTAGSELGDGGQSEN 288
Db      235 PNGSGDGGKGGNG-----GAGGN-----GGSIGANSQ-IVGSGGAGGAGGAGNGSLSSG 285
Qy      289 QKPLFILEDGASLKNVTMGDDGADGIHLYGDAKIDNLHVTNVG 331
Db      286 E-----GKGGGDGH-GGDGVGNGSSVTQGG 310

RESULT 17
B70807
hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70807
R: Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
  Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
  Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.;
  Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: B70807
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-1079 <COL>
A: Cross-references: GB:AL020222; GB:AL123456; NID: g3261554; PIDN: CAAL17749.1; PID: g2924444
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: Rv3512
C: Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 7.8%; Score 180.5; DB 2; Length 1079;
Best Local Similarity 25.3%; Pred. No. 0.019;
Matches 81; Conservative 17; Mismatches 105; Indels 117; Gaps 15;

Qy      7 NNNTSSPGFLFQSGDNGLGHNNANSALGQPIDRQTTEQMAQLLAELKLSLSPSGNA 66
Db      700 NCGAGTGSQ---NGNGSGGNGGNGAGNG-----GNS 728
Qy      67 AT-----GAGGNDQTTGVNAGLNGRKGTTAGTTTPQ-----SDSQNMLSEWNGNGLDQAIT 117
Db      729 GTGSGDGGAGNGGAAGTGTGGDGLTGTGTGSGGGTGGDGGN-----GGNGADNTAN 793
Qy      118 PDGQGGGQIGDNPILLKAMLKLIARMMDGQSDQFGQPTGTGNNASASSCTSSSGSPFNDLSG 177
Db      784 MTAQAGDGGNG-----GDGFGGAGAGGGGLTAGANGTGG-----QG 822
Qy      178 GKAPSGNSPGSNYPVSTPSPSTPTSPSPPLDPSPSPPTKAAGGSTPVTVDHPDVPVSAGI 237
Db      823 GAGDGG---GN-----GATGGHGLPTD--DPCGNGGT 849
Qy      238 GAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTTAGSEL-----GDGGQSENKQPLF 293
Db      850 G-GNGGTGGTGAGIGSLGGGTGGDGGN--GGNGGTGGEGEVGGAGGTGGAGNG----- 902
Qy      294 ILEDGASLKNVTMGDDGADG 313
Db      903 --GDGG-----TGGTGGGDG 915

RESULT 18

```


F96559
 Hypothetical protein F5F19.6 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: F96559
 R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C:Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96559
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-730 <STO>
 A:Cross-references: UNIPROT:Q9ZU23; GB:AR005173; NID:g4220464; PIDN:AAD12691.1; GSPDB:CN
 C:Genetics:
 A:Gene: F5F19.6
 A:Map position: 1

Query Match 7.7%; Score 179; DB 2; Length 730;
Best Local Similarity 23.2%; Pred. NO. 0.015;
Matches 115; Conservative 52; Mismatches 158; Indels 170; Gaps 24;

QY	9	NTSSPGLFQSGDNGNLGHNNANSALGOOPIDROTIEQMAQLLELLKSLSPQSGNAAT	68
Db	132	STQKPGASGIGSDSGSGIG-----SAGTNPAGDGTRE-----TEKNAGSKTSSSAGT	180
QY	69	GAG-----GNDQTGTGVGNAGGLNGRKGTTAGTTPQSDSQNMLSEGMNNGLDQAIFDQGGG	124
Db	181	NFGASAVNGETE-----KNAGGSKPSSGSAGTNPGA-----SAGANGETKNGVGSKPSSG	232

Qy 125 QIGNPLLKAMLKLIARNMDGSDQFGQPCT---GNNSASGT-----164

233 KAGTNP-----GANAGGNGTEKNAGGSKSSGSGARTNPGASAGNGTGV 277

Db 278 SNIGDTESNAGGKSKNDGANNGASGIESNAG--SCTGTFAGAGGTGGTGIDTESIAGGSKTN 335

QY 218 AAGSTPVTDHPDPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQVFDGKGQTFYAG 277

DB	336	SGNGGT--ND----	---GASGIGSNDGSTGTPGAG-----	GGTDSNIEG	371
QY	278	SELGDGGGSGENOKPI,FTI,ENGAS-I,KNTWTMCDGACDCTIIV	WQDQVAVTAT	ITTTT	282

Db 372 TENNVGGKETN-----PGASGIGN---SDSGTGTSPG-----TESNADG 408

QY 337 VKENSAGKSHVEITNSSFEHASDKILQLNADTNLSVDNWKAKDFGTFVRTNGGQQGN-W 395

DB	409	TKTNTGGKESN	--TGS-----	ESNTNSSPQKLEAQ	-----GGNGGNQW	444
QY	396	DIALISH	-----	ISAENGKKEFWKSN	-----SEC	400

445 DDGTDHGDYMKIHVAVGGLGIEQIRFDYVKNQGLKEGPHGVKGRGTSTTEISHPDEYL 504

QY 429 GDVENHYKVPMSANL 443

505 VSVEGLYD---SSNI 516

RESULT 19
HH70846

Mycobacterium geymurei
 #Species: Mycobacterium tuberculosis
 #Date: 17-Jul-1998 #sequence revision 17-Jul-1998 #text change 20-Jun-2000
 Mycobacterium geymurei protein KV3345C - Mycobacterium tuberculosis (strain H37Rv)

Accession: H70846
R. Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Sqaers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: H70846
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1538 <COL>
 A;Cross-references: GB:AL021841; GB:AL123456; NID:G3261517; PIDN:CAAI7117.1; PID:G3261518
 A;Experimental source: strain H37Rv
 C;Genetics:
 A;Gene: RV3345c
 C;Superfamily: collagen alpha 1(IV) chain

Query Match	7.7%;	Score 179;	DB 2;	Length 1538;
Best Local Similarity	24.3%;	Pred. No. 0.036;		

QY 19 SGGDNLGHHNANGALGOOPIDRQTIEQMAQLLAELLKSLSPQSGNA-----ATGAGGND 74
 :||| ||| |
 Db 195 NGGAGGAGTGGAG-----GNWLFGAGGGGAG 226

Qy 75 QTTGVGNAGGL--NGRKGTA^{CT}TPQSDSQNMLSENGNGLDQA^{IT}PDGQGGGQIGDNPL 131

DB 227 TNGGVGSGGFVYNGGAGGIGG-----IGGIGNGGDAGLFGNGGAGGAGA--- 273

QY 132 LKAMLKLIARMWMDGSDNDFGPGCTGNN-----GAGCGCGCGCGCGG 137

[illegible]

QY 171 PFNDLGGKAPSGNSP-----SGNYSPVSTESPSTPTS-----PTSPLDF 211

DB	332	AVMNGAGENGCHGNPNPVGVGAGGACGILLAGHGAAGATFTSGNGDGGIGATANSPLQ-	390
QY	212	PSSPTKAAGGSTPVTDDHDPDPVGSAGI--GACNLSVAETSACANOMTPTDNTNTNTGCGTSSG	288

[illegible]

271 GQTFTAGSELGGGQSENQKPLFILEDGASLKNVTMGDDGADGTHLYGDAKIDNLHVTNV 330

435 GTNCGAGGCGGCGGCAQHG----DGGVGGKGGAGSGGAG-----N 474

db 475 GFDAATLGSPGADGGMGNGGKGGDGKAGDGGAGAGDVTLAVNQAGGDDG-----NG 529

390 GQQG 393

530 GEVG 533

RESULT 20
877300

Protein: 6803; Accession: P05000; Gene: *syn*; Species: *Synechocystis* sp. PCC 6803; Strain: PCC 6803; Variety: PCC 6803; Synonym: Synechocystis sp. (strain PCC 6803)

```
#sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
#date: 25-Apr-1997
#accession: S77300
S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Minoda, S.; Kaneko, T.; Sato, T.
```

Ogino, K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M. J. J. Nat. Res. 3, 109-136, 1996

Sequence analysis of the genome of the unicellular cyanobacterium *Synechococcus* sp. S74322; Reference number: S74322; MUID:97061201; PMID:8905231

```

;Accession: S77300
;Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA

```

;Residues: 1-3016 <KAN>
;Cross-references: UNIPROT:P73590; EMBL:D90907; GB:AB001339; NID:G1652618; PIDD:
;Note: The nucleotide sequence information was derived from the complementary DNA
sequence.

19880000: the macrolide sequence was submitted to the EMBL Data Library, June 1988

Query Match 7.7%; Score 179; DB 2; Length 3016;
Best Local Similarity 24.5%; Pred. No. 0.082;
Matches 114; Conservative 34; Mismatches 153; Indels 164; Gaps 23;

Qy	1	MSILTLNNTSSPGLFQSGD---	NGLGHNANSALGQQPIDRQTEQMAQLLAELL--	55
Db	2476	LSYMLQNG-----YAKGGDVSGGGGAGAGGAL-----	TLQGSTVIIDNVSP	2520
Qy	56	-----KSLLSPOSGNAATGAGNDQTTGVGNAGGLNGRKG--	TAGT	94
Db	2521	DSNRVGGNGTAGARGGGGIFGTDPDRPSAGGGGNFNAGSNANNNGAAGSVGDDKAGG	2580	
Qy	95	TPQSDSQNMLSEMNGLDQAITPDGOGGQIGDNPLLKAWLKLIARMMD-----	QQ	146
Db	2581	TGGTGGTGGFGIGGGG-----GGGGSSSGSP-----	FDTPRNGSGG	2620
Qy	147	SDQFGQPGTGNNSASGTSFGSGSPFNDLSGKAPSGNSPSGNSPVSTPSTPTPT	206	
Db	2621	PGGAGGFGAGGAGGGGGGGGQNPQDDKRGFGGSGSGGN--PI-----	2665	
Qy	207	SPLDFPSSPTKAAGGSTPTVTDHPDVPVGSAGIGAGNSVAFTSAGANQTVLHDTITVKAGQV	266	
Db	2666	-----CGN-----GSVGSQG-CAGN-----	TVEGGS-	2685
Qy	267	FDGKGQFTTAGSELGDCGQSENQKPLFILEDGASLXNVTMGDDGADGCIHLYGDAKI--	DN	324
Db	2686	-PGNGSGGGGAGLG-GALFINQATVITNSQFS-GNTTQGTGCGNSGQALGGAIFITDN	2742	
Qy	325	LHVTNVGEDAITVKPNSAGKKSHEITNS--SFEHSDKILQL-----	NADTNL-----	371
Db	2743	SFTVGAG---LTFSNSAPSPSGSFSNAYQFQNNNDVYGTIDSLSNASTNLFVLPPLT	2799	
Qy	372	-SVDN---VKAKDFGTFTVINGGQ-----	GNWDL	397
Db	2800	PEVQNEISVVAKDNLPFVQLPDGSEVSLYEDQPFVSGSFGNQI	2844	

Search completed: January 25, 2005, 12:08:03
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 11:56:15 ; Search time 64 Seconds
(without alignments)
4018.632 Million cell updates/sec

Title: US-09-879-248-6
Perfect score: 2310
Sequence: 1 MSLLTNNNTSSPGLFQSG.....LGDVENHYKVPMSANLKVAE 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Uniprot_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2310	100.0	447	2	O54508 erwinia amy
2	2277	98.6	447	2	O9LAW2
3	1916.5	83.0	450	2	O6XDB8
4	1916.5	83.0	450	2	AAQ17046
5	1916.5	83.0	450	2	AAQ17046
6	1005	43.5	479	2	O6RKS2
7	1005	43.5	479	2	O6RKS2
8	1005	43.5	479	2	O6RKS2
9	705	30.5	138	2	O79AW7
10	589	25.5	424	2	O8RP12
11	567.5	24.6	441	2	O87327
12	559	24.2	424	2	O87264
13	559	24.2	424	2	O7C415
14	472.5	20.5	386	2	O4A9P3
15	423.5	18.3	307	2	O8KKZ2
16	415.5	18.0	380	2	O8XVQ5
17	362.5	15.7	224	2	O9RHM0
18	354	15.3	303	2	O8PHI5
19	353.5	15.3	221	2	O34310
20	349.5	15.1	222	2	O9X622
21	346.5	15.0	324	2	O8PBA4
22	341	14.8	222	2	O8LOR5
23	256	11.1	276	2	O829M8
24	226.5	9.8	266	2	O9EX16
25	216.5	9.4	215	2	O00843
26	207	9.0	1660	2	O79FD4
27	207	9.0	1660	2	CAE55496
28	207	9.0	1665	2	O7D721
29	206	8.9	222	2	O7S4B7
30	199.5	8.6	1306	2	O7D9L6
31	199.5	8.6	1306	2	O7UIQ7

ALIGNMENTS

32	199.5	8.6	1306	2	O6MX28
33	199.5	8.6	1306	2	CAE55300
34	194.5	8.4	233	2	Q00845
35	194	8.4	1928	2	O8T9H1
36	193	8.4	240	2	O93877
37	193	8.4	694	2	O7D724
38	193	8.4	694	2	CAE55495
39	191	8.3	555	2	O8SX80
40	191	8.3	929	2	O9NBL3
41	191	8.3	929	2	O9NBW0
42	191	8.3	929	2	AAE64813
43	191	8.3	939	2	O7KRH9
44	191	8.3	939	2	O9NHQ0
45	191	8.3	939	2	AAE64812
46	191	8.3	2280	2	O9V8E6
47	191	8.3	2302	2	O9N693
48	190.5	8.2	1360	2	O7TWC4
49	190.5	8.2	1382	2	O6BN56
50	190.5	8.2	1382	2	O7TWC0
51	189.5	8.2	628	2	O8VJ19
52	188	8.1	540	2	O7T2I3
53	188	8.1	639	2	O79FJ9
54	188	8.1	639	2	CAE55430
55	188	8.1	650	2	O8VJW1
56	187.5	8.1	622	2	O8VKN3
57	187	8.1	795	2	O7U020
58	186.5	8.1	591	2	O6MX50
59	186.5	8.1	591	2	CAE55268
60	186.5	8.1	606	2	O7U2C0
61	186	8.1	929	2	O9NGW5
62	186	8.1	1329	2	O79FP2
63	186	8.1	1329	2	CAE55390
64	186	8.1	1408	2	O7U022
65	186	8.1	1408	2	O8VK17
66	185.5	8.0	809	2	P90534
67	185.5	8.0	1381	2	O6MWW9
68	185.5	8.0	1381	2	CAE55603
69	185.5	8.0	1489	2	O6MWW6
70	185.5	8.0	1489	2	CAE55607
71	185	8.0	2310	2	O9GRA9
72	184.5	8.0	1150	2	O7TYG8
73	184	8.0	987	2	O89CB5
74	184	8.0	1384	2	O8VIZ1
75	183.5	7.9	242	2	O4U701
76	183	7.9	588	2	O6MXX7
77	183	7.9	588	2	CAE55590
78	183	7.9	775	2	O8VJ15
79	183	7.9	1507	2	O8VJ23
80	182.5	7.9	244	2	Q00851
81	182.5	7.9	1217	2	O8VIY9
82	182	7.9	606	2	O79FV3
83	182	7.9	606	2	CAE55332
84	182	7.9	609	2	O8VKC5
85	182	7.9	731	2	O6MXX5
86	182	7.9	731	2	CAE55593
87	182	7.9	1991	2	O7RTD0
88	181.5	7.9	484	2	O6MXX1
89	181.5	7.9	484	2	CAE55586
90	181.5	7.9	1618	2	O6FQ10
91	181	7.8	749	2	O7TMC6
92	181	7.8	1376	2	O7SSH8
93	181	7.8	3056	2	O7USQ0
94	180.5	7.8	1079	2	O6MWW7
95	180.5	7.8	1079	2	CAE55606
96	180.5	7.8	1715	2	O8VIZ0
97	180	7.8	626	2	O7TWM2
98	179.5	7.8	588	2	O9QEK6
99	179.5	7.8	608	2	O7U125
100	179.5	7.8	717	2	O75FLO

O6mx28	mycobacteri
Caes5300	mycobacte
Q00845	nectria hae
O8T9H1	drosofila
O93877	fusarium ox
O7D724	mycobacteri
Caes5495	mycobacte
Q8SX80	drosofila
Q9NBL3	drosofila
Q9NBW0	drosofila
AAE64813	drosofila
O7KRH9	drosofila
O9NHQ0	drosofila
AAE64812	drosophil
O9V8E6	drosofila
O9N693	drosofila
O7TWC4	mycobacteri
O6BN56	debaromyce
O7TWC0	mycobacteri
O8VJ19	mycobacteri
O7T2I3	mycobacteri
O79FJ9	mycobacteri
Caes5430	mycobacte
O8VJW1	mycobacte
O8VKN3	mycobacteri
O7U020	mycobacteri
O6MX50	mycobacteri
Caes5268	mycobacte
O7U2C0	mycobacteri
O9NGW5	drosofila
O79FP2	mycobacteri
Caes5390	mycobacte
O7U022	mycobacteri
O8VK17	mycobacteri
P90534	dictyosteli
O6MWW9	mycobacteri
Caes5603	mycobacte
O6MWW6	mycobacteri
Caes5607	mycobacte
O9GRA9	drosofila
O7TYG8	mycobacteri
O89CB5	bradyrhizob
O8VIZ1	mycobacteri
O4U701	nectria hae
O6MXX7	mycobacteri
Caes5590	mycobacte
O8VJ15	mycobacteri
O8VJ23	mycobacteri
Q00851	nectria hae
O8VIY9	mycobacteri
O79FV3	mycobacteri
Caes5332	mycobacte
O8VKC5	mycobacteri
O6MXX5	mycobacteri
Caes5593	mycobacte
O7RTD0	plasmodium
O6MXX1	mycobacteri
Caes5586	mycobacte
O6FQ10	candida gla
O7TMC6	mycobacteri
O7SSH8	neutrospora
O7USQ0	rhodospirell
O6MWW7	mycobacteri
Caes5606	mycobacte
O8VIZ0	mycobacteri
O7TWM2	mycobacteri
O9QEK6	cynomolgue
O7U125	mycobacteri
O75FLO	leptospira

```
RESULT 1
OS4508 ID O54508 PRELIMINARY; PRT; 447 AA.
AC O54508;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE HrpW protein.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFBP1430;
RX MEDLINE=98086111; PubMed=9426142;
RA Gaudriault S., Malandrin L., Paulin J.P., Barny M.A.;
RT "DepA, an essential pathogenicity factor of Erwinia amylovora showing
RT homology with AvrE of Pseudomonas syringae, is secreted via the Hrp
RT secretion pathway in a DspB-dependent way.";
RL Mol. Microbiol. 26:1057-1069(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CFBP1430;
RX MEDLINE=98316710; PubMed=9654138;
RA Gaudriault S., Brisset M.N., Barny M.A.;
RT "HrpW of Erwinia amylovora, a new Hrp-secreted protein.";
RL FEBS Lett. 428:224-228(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba321;
RA Kim J.F., Zumoff C.H., Beer S.V.;
RT "HrpW, a new harpin of Erwinia amylovora, is a member of a family of
RT pectate lyases.";
RL Phytopathology 87:0-0(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba321;
RX MEDLINE=98422475; PubMed=9748455;
RA Kim J.F., Beer S.V.;
RT "HrpW of Erwinia amylovora, a new harpin that contains a domain
RT homologous to pectate lyases of a distinct class.";
RL J. Bacteriol. 180:5203-5210(1998).
DR EMBL; Y13831; CAA74158.1; -.
DR EMBL; U94513; AAC62314.1; -.
DR PIR; T18447; T18447.
DR HSP; Q9RH0; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Query Match 100.0%; Score 2310; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.3e-110;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDQRTIEMQAQLLAELLSLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDQRTIEMQAQLLAELLSLS 60

QY 61 PQSGNAATGAGGNDQTTGVGNAGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120
DB 61 PQSGNAATGAGGNDQTTGVGNAGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120

QY 121 QGGGQIGDNPILLKAMLKLIARMMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180
DB 121 QGGGQIGDNPILLKAMLKLIARMMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180

QY 181 PQSGNPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240
DB 181 PQSGNPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300

QY 301 LKNTVMGDDGADGIGHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHSD 360
DB 301 LKNTVMGDDGADGIGHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHSD 360

RESULT 2
Q9LAW2 ID Q9LAW2 PRELIMINARY; PRT; 447 AA.
AC Q9LAW2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Harpin HrpW.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSP; Q9RH0; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45340 MW; 0BBAAE3871EDC2F6 CRC64;

Query Match 98.6%; Score 2277; DB 2; Length 447;
Best Local Similarity 98.7%; Pred. No. 1.1e-108;
Matches 441; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDQRTIEMQAQLLAELLSLS 60
DB 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQOPIDQRTIEMQAQLLAELLSLS 60

QY 61 PQSGNAATGAGGNDQTTGVGNAGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120
DB 61 PQSGNAATGAGGNDQTTGVGNAGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120

QY 121 QGGGQIGDNPILLKAMLKLIARMMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180
DB 121 QGGGQIGDNPILLKAMLKLIARMMDGQSDQFGQPGTGNNSASSTSSGSGSPNDLSGGKA 180

QY 181 PQSGNPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240
DB 181 PQSGNPSGNYSPVSTFSPSTPTSPPLDFFSSPTKAAGGSTPTVDHDPVGSAGIGAG 240

QY 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300
DB 241 NSVAFTSAGANQTVLHDTITVKAGQVFDGKQOTFTAGSELGDGQGSSEKQKPLFILEDGAS 300

QY 301 LKNTVMGDDGADGIGHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHSD 360
DB 301 LKNTVMGDDGADGIGHLYGDAKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSFEHSD 360
```

```
Db 301 LKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVEITNSPEHASD 360
QY 361 KILQLNADTNLSVDNVKAKDFGTFTVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSDSEGLN 420
Db 361 KILQLNADTNLSVDNVKAKDFGTFTVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSDSEGLK 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
Db 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 3
Q6XDB8 PRELIMINARY; PRT; 450 AA.
ID AAQ17046;
AC AAQ17046;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN HrpW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RL and expression of hrpP encoding elicitor of the hypersensitive
RL response.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY237642; AAQ17046.1; -
DR InterPro: IPR004898; Pectate lyase; 1.
DR Pfam: PF03211; Pectate lyase; 1.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 83.0%; Score 1916.5; DB 2; Length 450;
Best Local Similarity 83.2%; Pred. No. 2.9e-90;
Matches 376; Conservative 23; Mismatches 46; Indels 7; Gaps 2;

QY 1 MSILTLNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQWQAQLLAELLKLLS 60
Db 1 MSVLTNISIPSSQGLFKPEDNGLSGQNTNSAQGHPIDRQTIQWQAQLLGELLKPLLS 60
QY 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120
Db 61 PQADNAA--AGSNDQTNVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 118
QY 121 QGGQIGDNPGLLKAMLLKLIARMMDGQSDQFGQPGTGNNSSAGTSSGSGSPFNDSL--- 176
Db 119 QGGQIGDNPGLLKAMLLKLIARMMDGQSDQFGQPGTGNNSSAGTSSGSGSPFNDSL--- 178
QY 177 -GKAPSGNSPNSPVSTFSPSTPTSPSLDPPSSPTKAAGSTPTVTDHPDPVGS 235
Db 179 LGSSSLGKASSGSGTPTNSFSPSTPTSPSLDPPSSPTKAAGSTPTVTDHPDPVGS 238
QY 236 GIGAGNSVAFTSAGANOTVLHDTITVKAGVDFGKGTFTAGSELGDDGQSEKQPLFIL 295
Db 239 GVAGNSVGTSAAGNPTVLHDTIIIVKAGQEFPGKGTFTAGSELGDDGQSEKQPLFIL 298
QY 296 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSF 355
Db 299 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVDITNSSF 358
QY 356 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSF 355
Db 358 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVDITNSSF 358
QY 356 EHASDKILQLNADTNLSVDNVKAKDFGTFTVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 415
Db 359 EHASDKILQLNADTNLSVDNVKAKDFGTFTVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 418
QY 416 SEGNAVNTSDISLGDVENHYKVPMSANLKVAE 447
Db 419 SEGNAVNTSDISLGDVENHYKVPMSANLKVAK 450

RESULT 5
AAS45453 PRELIMINARY; PRT; 450 AA.
ID AAS45453
AC AAS45453;
```

```
Db 359 EHASDKILQLNADTNLSVDNVKAKDFGTFTVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 418
QY 416 SEGNAVNTSDISLGDVENHYKVPMSANLKVAE 447
Db 419 SEGNAVNTSDISLGDVENHYKVPMSANLKVAK 450

RESULT 4
AAQ17046 PRELIMINARY; PRT; 450 AA.
ID AAQ17046;
AC AAQ17046;
DT 03-MAR-2004 (TrEMBLrel. 27, Created)
DT 03-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 03-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN HrpW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RL and expression of hrpP encoding elicitor of the hypersensitive
RL response.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY237642; AAQ17046.1; -
DR InterPro: IPR004898; Pectate lyase; 1.
DR Pfam: PF03211; Pectate lyase; 1.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 83.0%; Score 1916.5; DB 2; Length 450;
Best Local Similarity 83.2%; Pred. No. 2.9e-90;
Matches 376; Conservative 23; Mismatches 46; Indels 7; Gaps 2;

QY 1 MSILTLNNTSSPGLFQSGDNGLGHNANSALGQOPIDRQTIQWQAQLLAELLKLLS 60
Db 1 MSVLTNISIPSSQGLFKPEDNGLSGQNTNSAQGHPIDRQTIQWQAQLLGELLKPLLS 60
QY 61 PQSGNAATGAGNDQTTGVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 120
Db 61 PQADNAA--AGSNDQTNVGNAGLNGRKGCTAGTTTQSDSQNNMLSEWNGNGLDQAITPDG 118
QY 121 QGGQIGDNPGLLKAMLLKLIARMMDGQSDQFGQPGTGNNSSAGTSSGSGSPFNDSL--- 176
Db 119 QGGQIGDNPGLLKAMLLKLIARMMDGQSDQFGQPGTGNNSSAGTSSGSGSPFNDSL--- 178
QY 177 -GKAPSGNSPNSPVSTFSPSTPTSPSLDPPSSPTKAAGSTPTVTDHPDPVGS 235
Db 179 LGSSSLGKASSGSGTPTNSFSPSTPTSPSLDPPSSPTKAAGSTPTVTDHPDPVGS 238
QY 236 GIGAGNSVAFTSAGANOTVLHDTITVKAGVDFGKGTFTAGSELGDDGQSEKQPLFIL 295
Db 239 GVAGNSVGTSAAGNPTVLHDTIIIVKAGQEFPGKGTFTAGSELGDDGQSEKQPLFIL 298
QY 296 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSF 355
Db 299 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVDITNSSF 358
QY 356 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVEITNSSF 355
Db 358 EDGASLKNVTMGDDGADGHIHYGDADKIDNLHVTNVGDAITVKPNSAGKSHVDITNSSF 358
QY 356 EHASDKILQLNADTNLSVDNVKAKDFGTFTVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 415
Db 359 EHASDKILQLNADTNLSVDNVKAKDFGTFTVRTNGGQGGNWDNLNLSHISAEDGKFSFVKSD 418
QY 416 SEGNAVNTSDISLGDVENHYKVPMSANLKVAE 447
Db 419 SEGNAVNTSDISLGDVENHYKVPMSANLKVAK 450

RESULT 5
AAS45453 PRELIMINARY; PRT; 450 AA.
ID AAS45453
AC AAS45453;
```


Best Local Similarity 49.9%; Pred. No. 1.1e-43;
Matches 231; Conservative 60; Mismatches 102; Indels 70; Gaps 15;

QY 8 NNTSSPGLFQSGDNGLGHNANSALGQPIDRQTEQMAQLLAELLSKLL-----59
Db 61 NTPSSDTG--SSQOAGI-GNGSSALGQN-----GSAADMLKLEALIPSKNGQEA 110
QY 60 -SPQSGNAATGAGN---DQTTGVNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQA 115
Db 111 GNPLSSGSSGAAGNGSGASPLTSSGAGGVGGAQ-----NPEDLSRSLLDQDSAGSALNNA 165
QY 116 ITPDGGGGGOIGDNPLLKAMKLIARMDGQSDQFGOPGTGNNSSASSTSSGSSPNDL 175
Db 166 INTADGGGGLSGNDLLKALLELIGNLMDSQKGEFGQP-----QSGGQSGGSP-----215
QY 176 SGKAPSGNSPSGNYSPVSTFSPPTSPSLDPPSSPTKAAG-----STPVTDHDPD 230
Db 216 STG-AQASSSGGGSPAA-----PSAPSSVGGNGGAASAPLTAAPT 257
QY 231 PV--GSAG-----IGAGNSVATPSAGANQTVLHDTITVKAQVFDGKGTFTAGSELGD 282
Db 258 GVDGSSAASPTASTAGAG-PVSPPTASANTVVDITIKVGPGEVFDGKGTFTASSKLG 316
QY 283 GGOSEKOKPLFILEDGASLKNNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSA 342
Db 317 GGOAEGQKPLFELAQOATLKNVVGDAADGVHVRGDAKIDNVHVTNVGDAITVKSNS- 375
QY 343 GKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFCTFVRTNGGQGNWDLNLSHI 402
Db 376 GKPAVEITNSSAQASDKIFQLNADANLTIDNFKAKDFCTFVRTNGGQGNWDLNLSNI 435
QY 403 SABDGKFSFKVSDSEGLNVNTSDISLGDVENHYKVPMSANLKV 445
Db 436 DAENGKFSFKVSDSEGLNVKGNINLTNNVNNHYKVPDSANLQV 478

RESULT 8
AAS20352 PRELIMINARY; PRT; 479 AA.
AC AAS20352;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DE HRPW.
GN HRPW.
OS Pectobacterium atrosepticum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1039;
RA Holvea M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,
RA Birch P.R.J., Toth I.K.;
RT "Rapid demonstration of a role early in disease development for the
RT type III secretion system of Erwinia carotovora subsp. atroseptica
RT SCRI1039 using a pooled transposon mutation grid."
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY96066; AAS20352.1; -. 3463E226CDF0406A CRC64;
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 43.5%; Score 1005; DB 2; Length 479;
Best Local Similarity 49.9%; Pred. No. 1.1e-43;
Matches 231; Conservative 60; Mismatches 102; Indels 70; Gaps 15;

QY 8 NNTSSPGLFQSGDNGLGHNANSALGQPIDRQTEQMAQLLAELLSKLL-----59
Db 61 NTPSSDTG--SSQOAGI-GNGSSALGQN-----GSAADMLKLEALIPSKNGQEA 110
QY 60 -SPQSGNAATGAGN---DQTTGVNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQA 115
Db 111 GNPLSSGSSGAAGNGSGASPLTSSGAGGVGGAQ-----NPEDLSRSLLDQDSAGSALNNA 165

QY 116 ITPDGGGGGOIGDNPLLKAMKLIARMDGQSDQFGOPGTGNNSSASSTSSGSSPNDL 175
Db 166 INTADGGGGLSGNDLLKALLELIGNLMDSQKGEFGQP-----QSGGQSGGSP-----215
QY 176 SGKAPSGNSPSGNYSPVSTFSPPTSPSLDPPSSPTKAAG-----STPVTDHDPD 230
Db 216 STG-AQASSSGGGSPAA-----PSAPSSVGGNGGAASAPLTAAPT 257
QY 231 PV--GSAG-----IGAGNSVATPSAGANQTVLHDTITVKAQVFDGKGTFTAGSELGD 282
Db 258 GVDGSSAASPTASTAGAG-PVSPPTASANTVVDITIKVGPGEVFDGKGTFTASSKLG 316
QY 283 GGOSEKOKPLFILEDGASLKNNVTMGDDGADGHIHYGDAKIDNLHVTNVGDAITVKPNSA 342
Db 317 GGOAEGQKPLFELAQOATLKNVVGDAADGVHVRGDAKIDNVHVTNVGDAITVKSNS- 375
QY 343 GKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFCTFVRTNGGQGNWDLNLSHI 402
Db 376 GKPAVEITNSSAQASDKIFQLNADANLTIDNFKAKDFCTFVRTNGGQGNWDLNLSNI 435
QY 403 SABDGKFSFKVSDSEGLNVNTSDISLGDVENHYKVPMSANLKV 445
Db 436 DAENGKFSFKVSDSEGLNVKGNINLTNNVNNHYKVPDSANLQV 478

RESULT 9
Q79AW7 PRELIMINARY; PRT; 138 AA.
AC Q79AW7;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE HRPW (Fragment).
GN Name=hrpw;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Es321;
RX MEDLINE=98115919; PubMed=9448330;
RA Bogdanove A.J., Kim J.F., Wei Z., Kolchinsky P., Charkowski A.O.,
RA Conlin A.K., Collmer A., Beer S.V.;
RT "Homology and functional similarity of an hrp-linked pathogenicity
RT locus, depeF, of Erwinia amylovora and the avirulence locus avrE of
RT Pseudomonas syringae pathovar tomato."
RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330(1998).
DR EMBL; U97504; AAC04849.1; -.
FT NON TER 138 138
SQ SEQUENCE 138 AA; 13788 MW; 4FE177177C74B3C6 CRC64;

Query Match 30.5%; Score 705; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 6e-29;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTEQMAQLLAELLSKLLS 60
Db 1 MSILTLNNNTSSPGLFQSGDNGLGHNANSALGQPIDRQTEQMAQLLAELLSKLLS 60
QY 61 POSGNAATGAGNDQTTGVNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120
Db 61 POSGNAATGAGNDQTTGVNAGGLNGRKGTTAGTTTQSDSQNMLSEMGNGLDQAITPDG 120
QY 121 QGGGOIGDNPLLKAMLKL 138
Db 121 QGGGOIGDNPLLKAMLKL 138

RESULT 10
Q8RP12 PRELIMINARY; PRT; 424 AA.
ID Q8RP12
AC Q8RP12;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Type III effector HrpPma.
 OS Pseudomonas syringae (pv. maculicola).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=59511;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ES4326;
 RX MEDLINE=21862332; PubMed=11872842;
 RA Guttman D.S., Vinatzer B.A., Sarkar S.F., Ranall M.V., Kettler G.,
 RA Greenberg J.T.;
 RT "A Functional Screen for the Type III (Hrp) Secretome of the Plant
 RT Pathogen Pseudomonas syringae.";
 RL Science 295:1722-1726 (2002).
 DR EMBL; AF458044; AAL84244.1; --
 DR HSSP; Q9RH00; 1EE6.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0030570; F:pectate lyase activity; IEA.
 DR InterPro; IPR011050; Pectin_lyase_like.
 DR InterPro; IPR004898; Pect_lyase.
 DR Pfam; PF03211; Pectate lyase; 1.
 DR Pfam; PF03211; Pectate lyase; 1.
 SQ SEQUENCE 424 AA; 42375 MW; 9C9B7E9E81E3886F CRC64;
 Query Match 25.58; Score 589; DB 2; Length 424;
 Best Local Similarity 38.08; Pred. No. 1.8e-22;
 Matches 156; Conservative 55; Mismatches 150; Indels 50; Gaps 11;
 QY 57 SLSPQSGNAATGAGGNDQTTGVNAGLNGKRTAGTTTPOSDSQNMLSEGNGLDQAI 116
 DB 20 SALSGLKSPQTDITLAGQNNQ-----AVDAEALLFGSDKQDVV-----FGTPENT 63
 QY 117 TPQGGGGGQIGD-----NPLKMKMLKLIARMWQGSDFQGPQGTGNNSSAGSTSS 167
 DB 64 APNLQNSQTGSSQSNLVKLFALIMSLQMLANTMKNQ-DTADPENWQDPFQNGNLG 122
 QY 168 GGSFFNDLGGKAPGNSPSGNSPVSTPSPSTPTSPSTPLDTPSSPTKAAG-----220
 DB 123 GDTSLADGGSTPDATGGGDDTTSATGEGGGDTTSATDGGGNTTAAAGGGGGGTR 182
 QY 221 -----GSTP-VTHD---PPVGSAGIGAGNSVAFTSAGANTVLDHDTITVKAGVFDG 269
 DB 183 SATSDSGSTPSTVDNGTTPNKGSAPTGTG-PVQVPKASGDTVVVNTIKVAGETFDG 241
 QY 270 KQGTTAGSELGGQGSQNKPLFILEDGLSKNVTMGDDGADGHIHYG--DAK--IDML 325
 DB 242 QGKTFASNALGGQGSQKPLFELAEAGATLKNVNLQNEADGHIHYKAATDAKVTIDNL 301
 QY 326 HVTNVGEDAITVKNSAGKSHVEITNSPEHASDKIQLNADTNLSVDNVKAKDFGTFV 385
 DB 302 HAENVGEDLITVKGEGGAKVTNLDIKNSSAQGADDKTIQLNADTNLVNVDGFKATDFGTV 361
 QY 386 RTNGGQGN-WDLNLISHAEDGKFSVKSDSEGLNVNTSDISLGDVENHY 435
 DB 362 CTNGGKQFNDMSIELNGIDASHGKFKALVKSDSDLLKATGDIAMTDVRHAY 412
 RESULT 11
 ID 087327 PRELIMINARY; PRT; 441 AA.
 AC 087327;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DE Hrp type III secreted protein.
 OS Pseudomonas syringae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX MEDLINE=98422476; PubMed=9748456;
 RX Charkowski A.O., Alfano J.R., Preston G., Yuan J., He S.Y.,

RP SEQUENCE FROM N.A.
 RC STRAIN=B728a;
 RX MEDLINE=98422476; PubMed=9748456;
 RA Charkowski A.O., Alfano J.R., Preston G., Yuan J., He S.Y.,
 RA Collmer A.;
 RT "The Pseudomonas syringae pv. tomato HrpW protein has domains similar
 RT to harpins and pectate lyases and can elicit the plant hypersensitive
 RT response and bind to pectate.";
 RL J. Bacteriol. 180:5211-5217 (1998).
 DR EMBL; AF037983; AAC62530.1; --
 DR HSSP; Q9RH00; 1EE6.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0030570; F:pectate lyase activity; IEA.
 DR InterPro; IPR011050; Pectin_lyase_like.
 DR InterPro; IPR004898; Pect_lyase.
 DR Pfam; PF03211; Pectate lyase; 1.
 DR Pfam; PF03211; Pectate lyase; 1.
 SQ SEQUENCE 441 AA; 44714 MW; 7B3B3146E51897BB CRC64;
 Query Match 24.6%; Score 567.5; DB 2; Length 441;
 Best Local Similarity 36.4%; Pred. No. 2.4e-21;
 Matches 156; Conservative 46; Mismatches 133; Indels 93; Gaps 14;
 QY 36 QQPIDRQTI-----EQWAQLLAEALLKSLLS---POSNAATGAGGNDQTTGVNAGLNGR 88
 DB 67 QNPTSSAATDPQSNVVKLLSALVTSLLQMLMLNKKQDTGQDSNEWQDPFQNEGGLG-- 124
 QY 89 KGTAGTTPOSDSQNMLSEGNGLDQAITPDGQGGQIGDNPFLKMLKLIARMWQGS 148
 DB 125 -----TPSA-----EGSDGT-----QEA 138
 QY 149 QFGQPGTGNNASSCTSSSGSPFNDLGGKAPGNSPS---GNYSPVST---ESPSPS 200
 DB 139 SGDEEGGTTAATGDDGGGTSPTTEGDDG---GTSPTAEGDGGGYSVTGADGSGAPS 194
 QY 201 TPTSPSLDFFSPPTKAAGGSTPTVDHDPD-VGSAGIGAGNSVAFTSAGANQ-----TV 254
 DB 195 T-----EDGTGGGGSDCVTPQVTPQIANPGRNSGNTVSDTTGSLGSGSEVNV 243
 QY 255 LHDITITVKAGVFDGKQTTTAGSELGGQGSQNKPLFILEDGLSKNVTMGDDGADGI 314
 DB 244 VKDTIKVAGVFDGHHGATFTADKSMGTGQDDEHQKPLFELAEAGVILKNVNLGENEADI 303
 QY 315 HLYGDAK-----IDNLHVTNVGEDAITVKNSAGKSHVEITNSPEHASDKIQLNAD 368
 DB 304 HV--NAKNSQVITDINVHAQNVGEMITVKGEGAKVTNLTNSANGADDKVIQLNAD 361
 QY 369 TNLSDNVKAKDFGTFVNTNGQQ--GNWDLNLSHAEDGKFSVKSDSEGLNVNTSDIS 427
 DB 362 THLVGDFKATDFGTLVNTNGKQFDMSEVLNGVDATHGKFKALVKSDSDLLKATGDI 421
 QY 428 LGDVENHY 435
 DB 422 MTDVKHAY 429
 RESULT 12
 ID 087264 PRELIMINARY; PRT; 424 AA.
 AC 087264;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DE HrpW.
 OS Pseudomonas syringae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX MEDLINE=98422476; PubMed=9748456;
 RX Charkowski A.O., Alfano J.R., Preston G., Yuan J., He S.Y.,

Collmer A.;
 "The pseudomonas syringae pv. tomato HrpW protein has domains similar to harpins and pectate lyases and can elicit the plant hypersensitive response and bind to pectate.";
 J. Bacteriol. 180:5211-5217(1998).
 EMBL; AF005221; AAC62526.1; -;
 HSSP; QBRHW; 1EE6.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0030570; F:pectate lyase activity; IEA.
 DR InterPro; IPR011050; Pectin lyase like.
 DR InterPro; IPR004898; Pect lyase.
 DR Pfam; PF03211; Pectate lyase; 1.
 DR SEQUENCE 424 AA; 42910 MW; 7AF4ED059BE79D2E CRC64;
 SQ
 Query Match 24.2%; Score 559; DB 2; Length 424;
 Best Local Similarity 36.1%; Pred. No. 6.1e-21;
 Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;
 QY 37 QPIDRQTIEQMAQLLAELKSL---LSPQSGNAATGAGNDQTTGVNAGLGRKGTAG 93
 DB 72 KPNDSQS--NIAKLISALIMSLQLMTNSNKKQDTNQEQPDSQAPFQNGGLG----- 122
 QY 94 TTPQSDSQNMLSEMGNGLDQAITPDQGGGQIGDNPPLKAMLKLIARWMDGSDQFGQP 153
 DB 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
 QY 154 GTGNNSASSTSSSGSGSPNDLSGGKAPSGNSGNYSPVSTPSPSTPTSPDLPFS 213
 DB 152 GGDTPATGGGGGGGTPTATGG--SGTPTATGGGGGVTPTITQL-----A 200
 QY 214 SPTKAAGSTPVTDPDPVGSAGIGAGNSVAFTSAGANQTVLHDTTVKAGQVFDGKGOT 273
 DB 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINVVKDTIKVGAGEVFDGHGAT 245
 QY 274 FTAGSELGGGQSGENOKPLFILEDGLSKNVNWTGDDGADGIHLG-----DAKIDNLHVTN 329
 DB 246 FTADKSMGNGDQGENOKPFLAEGATLKNVNGENEGVDGIHVAKNAQAEVTDINHAQN 305
 QY 330 VGEDAITVPNSAGKSHVEITNSSFEHASDKILQLNADTNLSVDNVKAKDFGTFVRTNG 389
 DB 306 VGEDLITVRGEGGAATVNLNKNSSAKGADDKVQLNANTHLKIDNFKADDFGTVRTNG 365
 QY 390 GQO-GNWDNLNLSHSAEDGKFSVKSDSEGLNVTNTSDISLGDVENVH 435
 DB 366 GKQFDDMSIELNGIANHGKFAVKSDSDDLKLTATGNIAVTMDVKHAY 412
 RESULT 13
 Q7C4I5 PRELIMINARY; PRT; 424 AA.
 AC Q7C4I5; Q7BM42;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Type III helper protein HrpW (Pto) (HrpW).
 GN Name=hrpW (Pto); Synonyms=hrpW; OrderedLocusNames=PSPT01373;
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_taxid=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
 RA Winn M.B., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
 RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
 RA Nelson W.C., Davidse T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
 RA Khouri H.M., Fedotova N.B., Tran B., Russell D., Berry K.J.,
 RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
 RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
 RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
 RA Bender C.L., White O., Fraser C.M., Collmer A.;

RT "The complete genome sequence of the Arabidopsis and tomato pathogen
 RT Pseudomonas syringae pv. tomato DC3000.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX MEDLINE=98422476; PubMed=9748456;
 RA Charkowski A.O., Alfano J.R., Preston G., Yuan J., He S.Y.,
 RA Collmer A.;
 RT "The Pseudomonas syringae pv. tomato HrpW protein has domains similar
 RT to harpins and pectate lyases and can elicit the plant hypersensitive
 RT response and bind to pectate.";
 RL J. Bacteriol. 180:5211-5217(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX MEDLINE=20243785; PubMed=10781092;
 RA Alfano J.R., Charkowski A.O., Deng W.L., Badel J.L.,
 RA Petnicki-Ocwieja T., van Dijk K., Collmer A.;
 RT "The Pseudomonas syringae Hrp pathogenicity island has a tripartite
 RT mosaic structure composed of a cluster of type III secretion genes
 RT bounded by exchangeable effector and conserved effector loci that
 RT contribute to parasitic fitness and pathogenicity in plants.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4856-4861(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX MEDLINE=22467693; PubMed=12580281;
 RA Fouts D.E., Badel J.L., Ramos A.R., Rapp R.A., Collmer A.;
 RT "A pseudomonas syringae pv. tomato DC3000 Hrp (Type III secretion)
 RT deletion mutant expressing the Hrp system of bean pathogen P. syringae
 RT pv. syringae 61 retains normal host specificity for tomato.";
 RL Mol. Plant Microbe Interact. 16:43-52(2003).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX Ramos A.R., Rehm A.H., Collmer A.R.;
 RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RX Ramos A.R., Rehm A.H., Collmer A.R.;
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AB016860; AA054895.1; -;
 DR EMBL; AF232004; AAF71503.1; -;
 DR InterPro; IPR011050; Pectin lyase like.
 DR InterPro; IPR004898; Pect lyase.
 DR Pfam; PF03211; Pectate lyase; 1.
 KW Complete proteome.
 SQ SEQUENCE 424 AA; 42910 MW; 7AF4ED059BE79D2E CRC64;
 Query Match 24.2%; Score 559; DB 2; Length 424;
 Best Local Similarity 36.1%; Pred. No. 6.1e-21;
 Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;
 QY 37 QPIDRQTIEQMAQLLAELKSL---LSPQSGNAATGAGNDQTTGVNAGLGRKGTAG 93
 DB 72 KPNDSQS--NIAKLISALIMSLQLMTNSNKKQDTNQEQPDSQAPFQNGGLG----- 122
 QY 94 TTPQSDSQNMLSEMGNGLDQAITPDQGGGQIGDNPPLKAMLKLIARWMDGSDQFGQP 153
 DB 123 -TPSADS-----GGGG-----TPDATGGGG-GDTP-----SATGGG 151
 QY 154 GTGNNSASSTSSSGSGSPNDLSGGKAPSGNSGNYSPVSTPSPSTPTSPDLPFS 213
 DB 152 GGDTPATGGGGGGGTPTATGG--SGTPTATGGGGGVTPTITQL-----A 200
 QY 214 SPTKAAGSTPVTDPDPVGSAGIGAGNSVAFTSAGANQTVLHDTTVKAGQVFDGKGOT 273
 DB 201 NPNRTSG-----TGSVSDTAGS-----TEQAGKINVVKDTIKVGAGEVFDGHGAT 245
 QY 274 FTAGSELGGGQSGENOKPLFILEDGLSKNVNWTGDDGADGIHLG-----DAKIDNLHVTN 329
 DB 246 FTADKSMGNGDQGENOKPFLAEGATLKNVNGENEGVDGIHVAKNAQAEVTDINHAQN 305
 QY 330 VGEDAITVPNSAGKSHVEITNSSFEHASDKILQLNADTNLSVDNVKAKDFGTFVRTNG 389
 DB 306 VGEDLITVRGEGGAATVNLNKNSSAKGADDKVQLNANTHLKIDNFKADDFGTVRTNG 365

```
Db 246 FTADKNGNGDQGENQKPMFELAGATLKNNVLGENEVDGIIHVAKNAQEVITINVHAQN 305
Qy 330 VGEDAITVKPNSAGKSHVEITNSPEHASDKILOINADTNLSVDNVKAKDPGFTFVRTNG 389
Db 306 VGEDLITVKGEGAAVTNINIKNSAGKADDDKVQVQNNANTHLKIDNFKADDFGTMVRTNG 365
Qy 390 GQO-GNWDNLNLSHSAEDCKPFSVKSDEGLNVNTSDISLGDEVNHY 435
Db 366 GQOFDDMSIELNGIEANHGKFAVKSDSDLLKATGNIAMTDVXHAY 412

RESULT 14
ID Q849P3 PRELIMINARY; PRT; 386 AA.
AC Q849P3;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Harpin.
GN Name=hrpW;
OS Pseudomonas syringae (pv. phaseolicola).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=319;
RN [1]
RP SEQUENCE FROM N.A.
RA Tsaltas D.;
RT "Biochemical, structural and molecular characterization of resistant
RT interactions between Pseudomonas syringae pv. phaseolicola and
RT phaseolus vulgaris.";
RL Thesis (2003), University of London Library, Senate House, Malet
RL Street, London, In press.
DR EMBL; AY210846; AAC50075.1; -.
DR HSSP; Q9RHW0; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 386 AA; 40459 MW; 1F079531EC895228 CRC64;

Query Match 20.5%; Score 472.5; DB 2; Length 386;
Best Local Similarity 32.1%; Pred. No. 1.4e-16;
Matches 141; Conservative 52; Mismatches 121; Indels 125; Gaps 17;

Qy 16 LFQSGDNLGL-GHNANSALGQPIDROT-----IEQMAQLLAELKSLSPQSGNA 66
Db 47 LFNSAAPKNVSPQPDNSV--QSPQDANNANPESNLIKFSALIVSLQLMLN-LNKQ 103
Qy 67 ATGAGGNDQTTGVGNAGLNGRKGTAGTTPOSQNNMLSEMGNNGLDQAITPDGQGGQI 126
Db 104 DTDQSSQWDPFQNGGLG-----TFTDS-----GDSGDSTLAATGDCG--- 145
Qy 127 GNPILLKAMKLILARMMDGSDQDFQPGCTGNSSASCTSSGSPFNDLSGKAPSGNSP 186
Db 146 GDTP-----TATGNN-----TPSVEGSSND--DGVTFQLAMP 176
Qy 187 SGNYSPTVSTPSPPTSPISPLDPPSSPTKAAAGSTPVTDHPDVPGSAGIGAGNSVAF 246
Db 177 N-----HTSGTGPVSD----- 187
Qy 247 SAGANQ-----TVLHDTITVKAGVFGKGTOTAGSELGDGQSENQKPLPILEDGASL 301
Db 188 TAGANDQAGKVIIVKDTIKVAANTVVDAGATFTADKSMGNGDQSENQKPLFELAKGATL 247
Qy 302 KNTVMGDGADGTHLYG-DAK-----INLHVTNVGEDAITVKPNSAGKSHVEITNSSPEH 357
Db 248 KNAHGENEVDGILHVAKDAQAQVTDINVAENVGDMITVKEGGAKVTNLIQKNSAEN 307
Qy 358 ASDKILQLNADTNLSVDNVKAKDPGFTFVRTNGQO-GNWDNLNLSHSAEDCKPFSVKS 416
Db 308 ADDKVFQLNANTHLNITGFENNFGLVTRNGEKQFDMDNKLKIDGVADNHGKFAVKSDS 367
```

```
Qy 417 EGLNVNTSDISLGDEVNHY 435
Db 368 E-----DLQLGQQHRH 379

RESULT 15
ID Q8KKZ2 PRELIMINARY; PRT; 307 AA.
AC Q8KKZ2;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Hypothetical protein hrpW.
GN Name=hrpW;
OS Rhizobium etli.
OG Plasmid symbiotic plasmid p42d.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CFN42;
RX MEDLINE=91193195; PubMed=2013564;
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
RT "Structural complexity of the symbiotic plasmid of Rhizobium
RT leguminosarum bv. phaseoli.";
RL J. Bacteriol. 173:2411-2419 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CFN42;
RX MEDLINE=97419521; PubMed=9274036;
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
RA Cevallos M.A., Davila G.;
RT "Sequence, localization and characteristics of the replicator region
RT of the symbiotic plasmid of Rhizobium etli.";
RL Microbiology 143:2825-2831 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=CFN42;
RX MEDLINE=22309397; PubMed=12421308;
RA Quintero V., Cevallos M.A., Davila G.;
RT "A site-specific recombinase (RinQ) is required to exert
RT incompatibility towards the symbiotic plasmid of Rhizobium etli.";
RL Mol. Microbiol. 46:1023-1032 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RA STRAIN=CFN42;
RX Ramirez M.A., Bustos P., Girard L., Rodriguez O., Cevallos M.A.,
RA Davila G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA STRAIN=CFN42;
RX Quintero V., Bustos P., Davila G.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA STRAIN=CFN42;
RX Gonzalez V., Bustos P., Medrano-Soto A., Ramirez-Romero M.A.,
RA Romero D., Salgado H., Hernandez-Gonzalez I., Hernandez-Celis J.C.,
RA Quintero V., Girard L., Rodriguez O., Flores M., Cevallos M.A.,
RA Collado-Vides J., Davilla G.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA STRAIN=CFN42;
RX Quintero V., Bustos P., Davila G.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80928; AA054988.1; -.
DR HSSP; Q9RHW0; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
```

```
DR InterPro: IPR004898; Pect_lyase.
DR Pfam: PF03211; Pectate_lyase; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 307 AA; 32912 MW; ED8844FA3AJA5E98 CRC64;

Query Match 18.3%; Score 423.5; DB 2; Length 307;
Best Local Similarity 37.7%; Pred. No. 3.6e-14;
Matches 121; Conservative 48; Mismatches 83; Indels 69; Gaps 14;

QY 144 DGSDQDQGPQTGNSSASGTS---SSGSGPFNDLSCGKAPSGNSPGNSPVSTSPPS 200
DB 22 DLESCLG---GSNRADTFTTIRTKVSPFEDF-----SGNPPT-----LTSYVPS 66
QY 201 TPTS-----PSPDLF-----PSPPTKAAGSTPTVDPDPVGSAGICAGN 241
DB 67 RETSENGMDSPPDKLLRKHINWQSDSKKVPDSDEKQATTLQT-TTEKPD----- 114
QY 242 SVAFTSAGANQTVLHDTITVKAGQVDFGKQGTFTAGSELGDDGQSENOKPLFILEDCASL 301
DB 115 -----LSKEGSIIVNEPIVVDGG-VFDGKGATYFASKLGGGQSQSETQSPFLFKGNAIL 169
QY 302 KNYTMGDGAGIHLGYDAKIDNLHVTNVGEDAITVKPNSAGKSKSHVEITNSSFEHASDK 361
DB 170 KNVDLGENGAGIHVYDGTATLENVNWQNVGEDALTVK---SAG---DITLIIGSAKGATDK 224
QY 362 ILQINADTNLSVNVKAKDFGTFTVTRNGGQGGWDLNLSHISABDGKFS-----FVKSDSE 417
DB 225 IFQINADTFYKDFVADGFTTLVTRNGKQIDAD-----VTIDGAFSHSGNVFRTDSS 279
QY 418 GLNVN-TSDISLGDVENHYKV 437
DB 280 LASVTFSLSDITLDDVRNWTRV 300

RESULT 16
Q8XVQ5 PRELIMINARY; PRT; 380 AA.
AC Q8XVQ5
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE PROBABLE HARPIN-RELATED PROTEIN.
GN Names-popW; Synonyms=RS00072; OrderedLocusNames=RS02775;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
DR EMBL; AL646071; CAD16482.1; -.
DR HSPB; Q8XVQ5; IEE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
KW Complete proteome.
SQ SEQUENCE 380 AA; 39809 MW; 06777197E97365D1 CRC64;

Query Match 18.0%; Score 415.5; DB 2; Length 380;
Best Local Similarity 30.5%; Pred. No. 1.2e-13;
Matches 143; Conservative 52; Mismatches 133; Indels 141; Gaps 20;

QY 8 NNTSSPGLFQSGDNGLGGHNANSALGQOPIDQRTIEOMAQLLAELLKSLSPQSGNAA 67
DB 9 NNHFQTPSTW-----NHDAGSS-----IDTSLQRAVQLDDQLVLOQL----- 45
QY 68 TGAGGNDQTTGVGNAGLNGRKGKTAGTTPOSQSNMLSEMG---NNGLDQAITDPDQGGGQ 125
DB 46 -----EARKLFG-----NMLNQPAGNAGQNH-----GGHGGGH 75
QY 126 IGDNPLLKAMLKUIARMMDGQSQDFQPGPT-GNNSASSGTSSSGSGSPFNDLSCGK-APSG 183
DB 76 -----HGSNGFGENGCRFGSPHANSAPQDLELPANKPNNGKHNTSA 117
QY 184 NSPSGNSPVSTPSPSTPSPLDFPSSPTKAAGGSTPTVDHPDPVGSAGICAGNSV 243
DB 118 STPTDQTAFST--SPTTGTSP--PSTSATGEKAYGVKVP-----PEPTGVVDVS----- 165
QY 244 AFTSAGANQTVLHDTITVKAGQVDFGKQGTFTAGSELGDDGQSENOKPLFILEDCASLKN 303
DB 166 -----KPIVVRKAGETFDGGKYRPTKEMGDSQNEHQKPLFLEPGATLKN 212
QY 304 VTMGDGAGIHLGYDAKIDNLHVTNVGEDAITVKPNSAGKSKSH----- 347
DB 213 VQY--SGDGIHLLGSAKLDRVNVNQVGEDAITI--DGAKNRAHDAKIAIGIDPASIPGGT 268
QY 348 --VEITNSSFEHASDKILOLADTNLSVNVKAKDFGTFTVTRNGG-----CQGNW 395
DB 269 PRVEITNSAPYGAOKLQINGDVIDQVKGMYNAGAKVFTRTNGGDTQIKATVNVQDSNF 328
QY 396 DLNLSHISABDGKFSF-----VKSDS--EGLNVNTSDISLGDVENHYK 436
DB 329 QNVSEAVFRDTSKFSTASPSDDVKSADPFGLAPDKSQVT-GTNKVSYK 376

RESULT 17
Q9RHW0 PRELIMINARY; PRT; 224 AA.
AC Q9RHW0
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Pectate lyase.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KSM-P15;
RX MEDLINE=2023260; PubMed=10759850;
RA Hatada Y., Saito K., Koike K., Yoshimatsu T., Ozawa T., Kobayashi T.,
RA Ito S.;
RT "Deduced amino-acid sequence and possible catalytic residues of a
RT novel pectate lyase from an alkaliphilic strain of Bacillus.";
RL Eur. J. Biochem. 267:2268-2275(2000).
DR EMBL; AB011839; BA87892.1; -.
DR PDB; 1ER6; X-ray; A=28-224.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
KW Lyase.
SQ SEQUENCE 224 AA; 23751 MW; 99D04821B09DE523 CRC64;

Query Match 15.7%; Score 362.5; DB 2; Length 224;
Best Local Similarity 46.1%; Pred. No. 3.3e-11;
Matches 83; Conservative 29; Mismatches 59; Indels 9; Gaps 4;

QY 248 AGANQTVLHDTITVKAGQVDFGKQGTFTAG-SELGDDGQSENOKPLFILEDCASLKNVTM 306
DB 25 AEAAPTIVVHETIRVPAGQTFDGKQTVVNPNTLGDGSAQENQKPIFRLAAGASLKNVVI 84
QY 307 GDDGAGIHLGYDAKIDNLHVTNVGEDAITVKPNSAGKSKSHVEITNSSFEHASDKILOLN 366
```

```
Db 85 GAPAADGVHCYGDCTTINWIVWEDALTLK-SSGT-----VNISGAAYKAYDKVQFIN 139
QY 367 ADTNLSVDNVKAKDGFVTRTNGGQGNWDLNLSHISAEDGKFSFKYKDSSE---GLNVT 423
Db 140 AAGTINIRNPADIDIGKLVQNGGTTVKVMVENCNISRVKDALRLTDSSTSGTIVNT 199

RESULT 18
OBPIH5
ID OBPIH5 PRELIMINARY; PRT; 303 AA.
AC OBPIH5
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE HrpW protein.
GN Name=hrpW; OrderedLocusNames=XAC2922;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202245; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan P., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katayama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Spinola L.A.P., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AB011934; AAC37767.1; .
DR HSSP; Q9RHW0; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR Pfam; PF004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
KW Complete proteome.
SQ
SEQUENCE 303 AA; 31949 MW; 765283B23852001B CRC64;
Query Match.
Best Local Similarity 15.3%; Score 354; DB 2; Length 303;
Matches 100; Conservative 45; Mismatches 85; Indels 96; Gaps 14;
QY 135 MKLIARMDQSDQFQCPGTG-----NNSASGTSSSGGSPNDLSGGKAPSGNSPGNY 190
Db 3 MSELLQRLI-----QTQFQFQNTSQVDMNPSHCTSNQG-----EGNEHGR- 46
QY 191 SPVSTFGPPSTPTSPDLPSPSPKAGG-----STPVT-DHPDP-----VGSAGIG 238
Db 47 -----RNGRGDNFLATTPTSDSHSDKSAHSTVTGSPVG 81
QY 239 AGNSVAFSTAGANOTVLTITVKAGQVFDGKGQFTTAGSELGDGQSGENKXPLFILEDG 298
Db 82 SGRITNVNS-----PIVVHKEGVDFGNHLYVGGSGIGDGSQSEHQPFVVEQG 131
QY 299 ASLKNVTMGDDGADGHLHYGDAKINLHVNTVNGDAITVKNPSAGKKSH----- 347
Db 132 GTLQVNM-SCGGDGHLLGDLATLKNVNLNVSEDAWTI--DGFGNEHDSRIAGTSPAG 198
PIR; A70045; A70045.
```

```
QY 348 -----VEITNSFPHASDKILQLNADTNLSVDNVKAKDGFVTRTNGGQGNWDLNLSH 401
Db 189 LPAPKPIELDSSFNADSKVIQDNOADVLLRNVSNGAKVKFRTNG---GHTDID-SH 244
QY 402 ISAEQDKFSFKV-----SDSEGLNVN 422
Db 245 VIVEDSTLKGKEAVFRTDAPGAHVS 270

RESULT 19
O34310
ID O34310 PRELIMINARY; PRT; 221 AA.
AC O34310; Q795F2;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE YvpA.
GN Name=yvpA; OrderedLocusNames=BSU34950;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98230327; PubMed=9570401;
RA Reizer J., Hoischen C., Tilgemyer F., Rivolta C., Rabus R.,
RA Stulke J., Karamata D., Sailer M.H. Jr., Hillen W.,
RA "A novel protein kinase that controls carbon catabolite repression in
RA bacteria."
RL Mol. Microbiol. 27:1157-1169 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lazarevic V., Soldo B., Rivolta C., Reynolds S., Maue C.,
RA Karamata D.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrali E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krohn J.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Li H., Masuda S., Maue C.,
RA Madigou C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambut R., Wedler E., Wedler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256 (1997).
DR EMBL; AF017113; AAC67291.1; .
DR EMBL; Z99121; CAB15500.1; .
DR PIR; A70045; A70045.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:04:36 ; Search time 20 Seconds
(without alignments)
1482.207 Million cell updates/sec

Title: US-09-879-248-6
Perfect score: 2310
Sequence: 1 MSILTLNNTSSPGLFQSG.....LGDVENHYKVPMSNLKVAE 447

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCUTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfilese1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2310	100.0	447	3	US-09-120-927-2
2	2310	100.0	447	4	US-09-431-614-6
3	559	24.2	424	3	US-09-120-817-2
4	559	24.2	424	4	US-09-431-614-14
5	359.5	15.6	197	3	US-09-402-668-2
6	341	14.8	221	3	US-09-198-956-4
7	341	14.8	221	4	US-09-670-141-4
8	190.5	8.2	62	3	US-09-402-668-10
9	181.5	7.9	2411	3	US-09-268-347-36
10	174	7.5	2042	4	US-09-077-098A-6
11	165.5	7.2	344	1	US-08-891-254-7
12	165.5	7.2	344	2	US-08-819-539-7
13	165.5	7.2	344	2	US-09-030-270A-7
14	165.5	7.2	344	3	US-08-984-207-7
15	165.5	7.2	344	3	US-09-013-587-7
16	165.5	7.2	344	4	US-09-086-118-27
17	165.5	7.2	344	4	US-09-431-614-15
18	165.5	7.2	344	5	PCT-US96-08819-7
19	160.5	6.9	907	4	US-09-010-928B-4
20	157.5	6.8	2870	4	US-09-479-467A-15
21	157.5	6.8	3178	4	US-09-479-467A-4
22	156.5	6.8	1912	1	US-08-409-995-4
23	156.5	6.8	1912	3	US-08-685-467-4
24	156	6.8	2039	4	US-09-077-098A-7
25	155.5	6.7	2353	3	US-09-377-155-33
26	155.5	6.7	2353	4	US-08-913-942-4
27	155.5	6.7	2353	3	US-09-669-974-33

28	155.5	6.7	2353	4	US-09-797-862-33	Sequence 33, Appl
29	155.5	6.7	2353	4	US-09-684-707-4	Sequence 4, Appl
30	155.5	6.7	2354	3	US-09-268-347-47	Sequence 47, Appl
31	155	6.7	385	5	PCT-US93-06243-2	Sequence 2, Appl
32	155	6.7	495	3	US-08-794-795-2	Sequence 2, Appl
33	155	6.7	495	3	US-09-249-200-2	Sequence 2, Appl
34	154.5	6.7	975	4	US-09-328-352-4764	Sequence 4764, Ap
35	154	6.7	571	3	US-09-134-001C-3865	Sequence 3865, Ap
36	151	6.5	385	1	US-08-891-254-3	Sequence 3, Appl
37	151	6.5	385	2	US-08-819-539-3	Sequence 3, Appl
38	151	6.5	385	5	PCT-US96-08819-3	Sequence 3, Appl
39	151	6.5	403	2	US-08-200-724A-2	Sequence 2, Appl
40	151	6.5	403	2	US-09-030-270A-3	Sequence 3, Appl
41	151	6.5	403	3	US-08-851-376A-2	Sequence 2, Appl
42	151	6.5	403	3	US-08-984-207-3	Sequence 3, Appl
43	151	6.5	403	3	US-09-013-587-3	Sequence 3, Appl
44	151	6.5	403	4	US-09-086-118-23	Sequence 23, Appl
45	151	6.5	403	4	US-09-431-614-3	Sequence 3, Appl
46	151	6.5	619	4	US-09-252-991A-26352	Sequence 26352, A
47	151	6.5	889	3	US-09-336-447A-15	Sequence 15, Appl
48	151	6.5	889	4	US-09-952-267B-15	Sequence 15, Appl
49	150.5	6.5	745	2	US-09-010-928B-28	Sequence 28, Appl
50	150.5	6.5	870	2	US-09-010-928B-2	Sequence 2, Appl
51	150	6.5	448	4	US-09-248-796A-19135	Sequence 19135, A
52	150	6.5	520	2	US-08-794-795-6	Sequence 6, Appl
53	150	6.5	520	2	US-09-249-200-6	Sequence 6, Appl
54	147.5	6.4	489	2	US-08-794-795-7	Sequence 7, Appl
55	147.5	6.4	489	3	US-09-249-200-7	Sequence 7, Appl
56	147.5	6.4	518	1	US-08-392-367B-2	Sequence 2, Appl
57	147.5	6.4	518	3	US-08-893-467A-2	Sequence 2, Appl
58	146.5	6.3	943	3	US-09-056-556-204	Sequence 204, App
59	146.5	6.3	943	4	US-09-072-596-199	Sequence 199, App
60	146.5	6.3	943	4	US-09-072-596-199	Sequence 199, App
61	146.5	6.3	943	4	US-09-072-596-199	Sequence 199, App
62	146.5	6.3	1180	4	US-09-206-942-65	Sequence 65, Appl
63	146.5	6.3	1180	4	US-09-206-942-65	Sequence 65, Appl
64	146.5	6.3	1776	4	US-09-556-877-179	Sequence 63, Appl
65	146.5	6.3	1776	4	US-09-556-877-179	Sequence 179, App
66	146.5	6.3	1776	4	US-09-556-877-179	Sequence 179, App
67	145	6.3	334	3	US-09-598-419-179	Sequence 179, App
68	145	6.3	334	3	US-09-060-756-728	Sequence 728, App
69	144.5	6.2	2736	4	US-09-670-314-728	Sequence 728, App
70	144	6.2	1581	3	US-09-252-991A-30227	Sequence 30227, A
71	143.5	6.2	679	3	US-09-110-517-2	Sequence 2, Appl
72	143.5	6.2	679	3	US-08-913-942-15	Sequence 15, Appl
73	143.5	6.2	679	3	US-09-268-347-26	Sequence 26, Appl
74	143.5	6.2	679	3	US-09-684-707-15	Sequence 15, Appl
75	141.5	6.1	541	4	US-09-268-347-30	Sequence 30, Appl
76	141.5	6.1	541	4	US-09-538-092-347	Sequence 347, App
77	140.5	6.1	215	4	US-09-543-681A-5434	Sequence 5434, Ap
78	140.5	6.1	412	4	US-09-248-796A-26345	Sequence 26345, A
79	140.5	6.1	1060	3	US-08-911-393-2	Sequence 2, Appl
80	140.5	6.1	1060	4	US-09-955-909-2	Sequence 2, Appl
81	138.5	6.0	1036	4	US-09-198-452A-580	Sequence 580, App
82	138.5	6.0	1036	4	US-09-206-942-73	Sequence 73, Appl
83	138.5	6.0	1475	4	US-09-538-092-1160	Sequence 1160, Ap
84	138	6.0	1477	4	US-09-206-942-71	Sequence 71, Appl
85	138	6.0	353	2	US-08-687-702-37	Sequence 37, Appl
86	137.5	6.0	2090	4	US-09-538-092-1081	Sequence 1081, Ap
87	137.5	6.0	481	4	US-09-248-796A-14281	Sequence 14281, A
88	137	5.9	300	3	US-09-724-623-77	Sequence 77, Appl
89	137	5.9	300	3	US-08-765-856-2	Sequence 2, Appl
90	137	5.9	300	3	US-08-935-009A-2	Sequence 2, Appl
91	136.5	5.9	842	5	PCT-US96-02331-15	Sequence 15, Appl
92	136.5	5.9	1368	4	US-09-270-767-45320	Sequence 45320, A
93	136.5	5.9	681	4	US-09-490-291-4	Sequence 4, Appl
94	136.5	5.9	681	4	US-09-490-291-6	Sequence 6, Appl
95	136.5	5.9	1477	1	US-08-038-682-4	Sequence 4, Appl
96	136.5	5.9	1477	2	US-08-302-832-4	Sequence 4, Appl
97	136.5	5.9	1477	2	US-08-530-198-4	Sequence 4, Appl
98	136.5	5.9	1477	2	US-08-469-880-4	Sequence 4, Appl
99	136.5	5.9	1477	2	US-08-728-470-4	Sequence 4, Appl
100	136.5	5.9	1477	2	US-08-617-697-4	Sequence 4, Appl
101	136.5	5.9	1477	3	US-08-719-641-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-120-927-2
; Sequence 2, Application US/09120927
; Patent No. 6262018
; GENERAL INFORMATION:
; APPLICANT: Kim, Jihyun Francis
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM
; TITLE OF INVENTION: ERWINIA AMYLOVORA AND ITS USE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hatgrave, Devans & Doyle LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,927
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/055,108
; FILING DATE: 06-AUG-1977
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1581
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-120-927-2

Query Match 100.0%; Score 2310; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.8e-175;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIEMQAQLLAELLS 60
DB 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIEMQAQLLAELLS 60
QY 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPILLKMLKLIARMMDGSDQFGPGTGNNSASSGTSSSGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPILLKMLKLIARMMDGSDQFGPGTGNNSASSGTSSSGSPFNDLSGGKA 180
QY 181 PSGNSPSGNSYSPVSTPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240
DB 181 PSGNSPSGNSYSPVSTPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVRKAGQVFDGKGQTTAGSELGQSGSENQKPLFILEDGS 300
DB 241 NSVAFTSAGANQTVLHDTITVRKAGQVFDGKGQTTAGSELGQSGSENQKPLFILEDGS 300
QY 361 KILQLNADTNLSVDNVKAKDFGT FVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGT FVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

only 2 domains
claim #1
3 domains
but compensating in large wdg & class patent
(possible 103)
\$ w/ 98 also disclosures
active fragment

QY 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVETITNSFEHSD 360
DB 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVETITNSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGT FVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGT FVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
RESULT 2
US-09-431-614-6
; Sequence 6, Application US/09431614
; Patent No. 6624139
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; FILE REFERENCE: 21829/41 (EBC-003)
; CURRENT APPLICATION NUMBER: US/09/431,614
; EARLIER FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/107,243
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-431-614-6

Query Match 100.0%; Score 2310; DB 4; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.8e-175;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIEMQAQLLAELLS 60
DB 1 MSILTLNNNTSSSPGLFQSGDNGLGHNANSALGQOPIDRTIEMQAQLLAELLS 60
QY 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120
DB 61 POSGNAATGAGGNDQTTGVGNAGGLNGRKTAGTTPOSQSNMLSEMGNGLDQAITPDG 120
QY 121 QGGGQIGDNPILLKMLKLIARMMDGSDQFGPGTGNNSASSGTSSSGSPFNDLSGGKA 180
DB 121 QGGGQIGDNPILLKMLKLIARMMDGSDQFGPGTGNNSASSGTSSSGSPFNDLSGGKA 180
QY 181 PSGNSPSGNSYSPVSTPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240
DB 181 PSGNSPSGNSYSPVSTPSTPTSPPLDPPSSPTKAAGGSTPVTHDPDVGAGIGAG 240
QY 241 NSVAFTSAGANQTVLHDTITVRKAGQVFDGKGQTTAGSELGQSGSENQKPLFILEDGS 300
DB 241 NSVAFTSAGANQTVLHDTITVRKAGQVFDGKGQTTAGSELGQSGSENQKPLFILEDGS 300
QY 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVETITNSFEHSD 360
DB 301 LKNVTMGDDGADGIHLGYGDAKIDNLHVTNVGDAITVKNSAGKSHVETITNSFEHSD 360
QY 361 KILQLNADTNLSVDNVKAKDFGT FVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
DB 361 KILQLNADTNLSVDNVKAKDFGT FVRTNGGQGNWDLNLSHISAEDGKFSFVKSDSEGLN 420
QY 421 VNTSDISLGDVENHYKVPMSANLKVAE 447
DB 421 VNTSDISLGDVENHYKVPMSANLKVAE 447

RESULT 3

US-09-120-817-2
; Sequence 2, Application US/09120817
; Patent No. 6172184
; GENERAL INFORMATION:
; APPLICANT: Collmer, Alan
; APPLICANT: Charkowski, Amy
; APPLICANT: Alfano, James R.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM
; TITLE OF INVENTION: PSEUDOMONAS SYRINGAE AND ITS USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,817
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/055,107
; FILING DATE: 06-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1741
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-120-817-2

Query Match 24.2%; Score 559; DB 3; Length 424;
Best Local Similarity 36.1%; Pred. No. 2.9e-36;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL---LSPOSNAATGAGNDQTTGVNAGLNGRKGATG 93
Db 72 KPNDQS--NIAKLISALIMSLQLMTNSNKKQDTNQEOPDSQAPFQNNGLG-----122
QY 94 TTPQSDSQNMLSEMGNNGLDQAITPGCGGGQIGDNPLLKAMLKLIARWMDGSDQFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151
QY 154 GTGNNSASGTSSTSGSGSPFNDLSGGKAPSGNSPGNSYSPVSTPSTPTSPSPLDFPS 213
Db 152 GGTPTATGGGGGGGTPTATGGG---SGTPTATGGGGGVTPOITPOL-----A 200

QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFPSAGANQTVLHDTITVKAGQVDFGKGT 273
Db 201 NPRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTMGDDGADGIHLG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNDQGENQKPMFELAEATLKNVNLGENEVDGIHVKAQNAQEVITDINVAQN 305
QY 330 VGEDAITVKNPSAGKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFGTFVRTNG 389
Db 306 VGEDLITVKGEGAAVTNINIKNSAKGADDKVQVQLNANTHLKIDNFKADDFGTFVRTNG 365

QY 390 GQO-GNWDNLNLSHISAEDCKFSFKVSDSEGLNVNTSDISLGDVYENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFKALVKSDDSLKLATGNIAMTDVKHAY 412

RESULT 5

US-09-402-668-2
; Sequence 2, Application US/09402668
; Patent No. 6172030
; GENERAL INFORMATION:
; APPLICANT: WADA, Yasunao
; APPLICANT: KASAI, Miyuki
; APPLICANT: SHIKATA, Shitsuw
; APPLICANT: SUZUMATSU, Atsushi
; APPLICANT: KOIKE, Kenzo
; APPLICANT: HATADA, Yuji
; APPLICANT: KOBAYASHI, Tohru

QY 390 GQO-GNWDNLNLSHISAEDCKFSFKVSDSEGLNVNTSDISLGDVYENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFKALVKSDDSLKLATGNIAMTDVKHAY 412

RESULT 4
US-09-431-614-14
; Sequence 14, Application US/09431614
; Patent No. 6624139
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/41 (SBC-003)
; CURRENT APPLICATION NUMBER: US/09/431,614
; CURRENT FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/107,243
; EARLIER FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
; US-09-431-614-14

Query Match 24.2%; Score 559; DB 4; Length 424;
Best Local Similarity 36.1%; Pred. No. 2.9e-36;
Matches 147; Conservative 54; Mismatches 132; Indels 74; Gaps 13;

QY 37 QPIDRTIEQMAQLLAELKSL---LSPOSNAATGAGNDQTTGVNAGLNGRKGATG 93
Db 72 KPNDQS--NIAKLISALIMSLQLMTNSNKKQDTNQEOPDSQAPFQNNGLG-----122
QY 94 TTPQSDSQNMLSEMGNNGLDQAITPGCGGGQIGDNPLLKAMLKLIARWMDGSDQFGQP 153
Db 123 -TPSADS-----GGGG-----TPDATGGG-GDTP-----SATGGG 151
QY 154 GTGNNSASGTSSTSGSGSPFNDLSGGKAPSGNSPGNSYSPVSTPSTPTSPSPLDFPS 213
Db 152 GGTPTATGGGGGGGTPTATGGG---SGTPTATGGGGGVTPOITPOL-----A 200

QY 214 SPTKAAGGSTPVTDHPDPVGSAGIGAGNSVAFPSAGANQTVLHDTITVKAGQVDFGKGT 273
Db 201 NPRTSG-----TGSVSDTAGS---TEQAGKINVVKDTIKVGAGEVDFGHGAT 245
QY 274 FTAGSELGCGQSENQKPLFILEDGASLKNVTMGDDGADGIHLG-----DAKIDNLHVTN 329
Db 246 FTADKSMGNDQGENQKPMFELAEATLKNVNLGENEVDGIHVKAQNAQEVITDINVAQN 305
QY 330 VGEDAITVKNPSAGKSHVEITNSSPEHASDKILQNLADTNLSVDNVKAKDFGTFVRTNG 389
Db 306 VGEDLITVKGEGAAVTNINIKNSAKGADDKVQVQLNANTHLKIDNFKADDFGTFVRTNG 365

QY 390 GQO-GNWDNLNLSHISAEDCKFSFKVSDSEGLNVNTSDISLGDVYENHY 435
Db 366 GKQFDDMSIELNGIEANHGKFKALVKSDDSLKLATGNIAMTDVKHAY 412

```
; APPLICANT: ITO, Susumu
; APPLICANT: TSUMADORI, Masaki
; TITLE OF INVENTION: Detergent Composition
; FILE REFERENCE: 2173-0116P
; CURRENT APPLICATION NUMBER: US/09/402,668
; CURRENT FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 9-091142 JAPAN
; PRIOR FILING DATE: 1997-04-09
; PRIOR APPLICATION NUMBER: 9-242736 JAPAN
; PRIOR FILING DATE: 1997-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/01613
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Bacillus sp.
; OTHER INFORMATION: Strain: KSM-P15
US-09-402-668-2

Query Match 15.6%; Score 359.5; DB 3; Length 197;
Best Local Similarity 46.3%; Pred. No. 6.6e-21;
Matches 81; Conservative 29; Mismatches 56; Indels 9; Gaps 4;

QY 253 TVLHDTITVKGQVFDGKQFTTAG-SELGDSQSENQKPLFILEDGASLKNVTMGDDGA 311
DB 3 TVVHETIRVPAGQTFDGGQTYVNPNTLGDGSAENQKPIFLEAGASLKNVVGAPAA 62

QY 312 DGIHLYGDAKIDNLHVTNCGEDAITVKNPSAGKSKSHVEITNSSFEHASDKILOLNADTNL 371
DB 63 DGVHCYGDCTINVTWEDVGEDALTK--SSGT---VNISGAAYKAYDKVQFINAAGTI 117

QY 372 SVDNVKADFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSE---GLNVNT 423
DB 118 NIRNFRADDIGKLRQNGGTYKVMNVNENCNISKVKDAILRTDSTSTGRIVNT 172

RESULT 6
US-09-198-956-4
; Sequence 4, Application US/09198956
; Patent No. 6165769
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schuelein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377-200-US
; CURRENT APPLICATION NUMBER: US/09/198,956
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 1344/97
; EARLIER FILING DATE: 1997-11-24
; EARLIER APPLICATION NUMBER: 60/067,240
; EARLIER FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-198-956-4

Query Match 14.8%; Score 341; DB 3; Length 221;
Best Local Similarity 41.7%; Pred. No. 2.3e-19;
Matches 73; Conservative 33; Mismatches 59; Indels 10; Gaps 2;

QY 254 VLHDTITVKGQVFDGKQFTTAGSELGDSQSENQKPLFILEDGASLKNVTMGDDGADG 313
DB 31 VVHKTIVVEKQTYDGKRLIAGPGLGDSQSDQKPIFKVEDGATLKNVVLGAPADG 90

QY 314 IHLYGDAKIDNLHVTNCGEDAITVKNPSAGKSKSHVEITNSSFEHASDKILOLNADTNL 373
DB 91 VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGSGARLAADKIFQINKASTFTV 145

QY 374 DNVKADFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLNVNTSDISL 428
DB 146 KNFTADQGGKFIQOLGGSTFKAVVNDICTITNNKEAIFRTDS-----STSSVTM 195

RESULT 7
US-09-670-141-4
; Sequence 4, Application US/09670141
; Patent No. 6429000
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schuelein, Martin
; APPLICANT: Lange, Niels Erik K.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Schnorr, Kirk
; TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
; FILE REFERENCE: 5377-200-US
; CURRENT APPLICATION NUMBER: US/09/670,141
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/198,956
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-670-141-4

QY 314 IHLYGDAKIDNLHVTNCGEDAITVKNPSAGKSKSHVEITNSSFEHASDKILOLNADTNL 373
DB 91 VHTYGNASINNVWEDVGEDALTVK-----SEGSVTINGSGARLAADKIFQINKASTFTV 145

QY 374 DNVKADFGTFVRTNGGQGNWDLNLSHISAEDGKFSFKVSDSEGLNVNTSDISL 428
DB 146 KNFTADQGGKFIQOLGGSTFKAVVNDICTITNNKEAIFRTDS-----STSSVTM 195

RESULT 8
US-09-402-668-10
; Sequence 10, Application US/09402668
; Patent No. 6172030
; GENERAL INFORMATION:
; APPLICANT: WADA, Yasunao
; APPLICANT: KASAI, Miyuki
; APPLICANT: SHIKATA, Shitsuw
; APPLICANT: SUZUMATSU, Atsushi
; APPLICANT: KOIKE, Kenzo
; APPLICANT: HATADA, Yuji
; APPLICANT: KOBAYASHI, Tohru
; APPLICANT: ITO, Susumu
; APPLICANT: TSUMADORI, Masaki
; TITLE OF INVENTION: Detergent Composition
; FILE REFERENCE: 2173-0116P
; CURRENT APPLICATION NUMBER: US/09/402,668
; CURRENT FILING DATE: 1998-10-08
```

```

; PRIOR APPLICATION NUMBER: 9-091142 JAPAN
; PRIOR FILING DATE: 1997-04-09
; PRIOR APPLICATION NUMBER: 9-242736 JAPAN
; PRIOR FILING DATE: 1997-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/01613
; PRIOR FILING DATE: 1998-04-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 10
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide from
; OTHER INFORMATION: primer
US-09-402-668-10

Query Match      8.2%; Score 190.5; DB 3; Length 62;
Best Local Similarity 65.6%; Pred. No. 3.3e-08;
Matches 39; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

QY 253 TVLHDTITVKAGVFDGKGTFTAG-SELGSGQSENQKPLFILEDGAALKNVMTMGDDGA 311
DB 3 TVVHETIRVPAGTDFGKGTGYVNPNTLGDGSAENQKPIFLEAGASLKNVVICAPAA 62

RESULT 9
US-09-268-347-36
; Sequence 36, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 36
; LENGTH: 2411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-36

Query Match      7.9%; Score 181.5; DB 3; Length 2411;
Best Local Similarity 20.6%; Pred. No. 2.6e-05;
Matches 126; Conservative 64; Mismatches 205; Indels 217; Gaps 26;

QY 4 LTUN-----NTSSPGLFQGGONGLGHN-----ANSALGOQPIDRQT 43
DB 1182 ITLGAAGAGTASNGNTISVTXDGISAGNKEITNVKSALKTKYKDTQNTAGATQPA-ANT 1240
QY 44 IEQMAQLLAELLSLSPQSGNAATGAGGNDQTTGVGNAGLNG-----RKGTAGTTPQ 97
DB 1241 AEVAKQDLVDLTK-----PATGAAGNADAKAPDTTAAVTGDLGKLGWLSAKKTADETQD 1296
QY 98 SD-----SONLSEMGNGLDQAITPDGQ-----GGQIGNPLLLKAMKLIIARM 142
DB 1297 KEFHAAVKANAEVEFVGKNGATVSAKTDNNGKHTVIDVAEKVGDG-----LEKD 1347
QY 143 MCGSQDQFGQGTGNN-----SASGTSSSGSPFNDLSGKAPSGNSPGNSPVSTFSP 198
DB 1348 TDGKIKLVNDTGNLLTVDTAKGASVA-----KGEFNAVTT---1385
QY 199 PSTPTSTPLDPSFSTKAAGSTPV-TDHPD--PVGSAGICAGNSVAP-----245
DB 1386 DATTAQGTNANERKVVVVGKSGNATATETDKKVVATVGVDAKAINDAATPVKVENDDSAT 1445
QY 246 -----TSAGANQTV-LHDTITTVKAGVVF-----DGKGTFTAGSE-----LGD 282
DB 1446 IDSPPTDGDANDALKAGDTITLAKGNLKVKGKGNITFALANDLSVKSATVSDKLSLGT 1505
QY 283 GGQSENQKPLFILEDGAAL-----KNVTMGDDGADGIHLVG-----DAKIDNLHVTNVGSD 333

```

```

DB 1506 NGKVN-----ITSDTKGLNFAKDSKTDGDA--NIHLNGIASLTITLNSGATTNLGNN 1558
QY 334 AIT-----VKPNSAGKSHVEITNSFEHASDKILOLNADTNL 371
DB 1559 GITDNEKKRAASVKDVLNAGMNVGVKGPASANN-----QVENIDFVATYDTVDFVSGDKOT 1614
QY 372 SVDNVKAKDPG-----TFVRTNGGQOQGN-- 394
DB 1615 TSVTVESKONGKRTVEYKIGAKTSVIKOHNGKLFTEGKELKADANNNGVTVTETDGRKDSGNGL 1674
QY 395 -----WDLNLSHISAEDGKFSFKVS-----DSEGLNVNTSDISLGDVEN 433
DB 1675 VTAKAVIDAVKAGWRVKTGTGANGQNDPATVASCNTVTFADGNGTTAEVTKANDGSITV 1734
QY 434 HYKVPMSANLKV 445
DB 1735 KYNVKVADGLKL 1746

RESULT 10
US-09-077-098A-6
; Sequence 6, Application US/09077098A
; Patent No. 6544519
; GENERAL INFORMATION:
; APPLICANT: TOKUNAGA, Eiji
;             SAKAGUCHI, Masaaki
;             MATSUO, Kazuo
;             HAMADA, Fukuaburo
;             TOKIYOSHI, Sachio
; TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/077,098A
; FILING DATE: 19-May-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/03222
; FILING DATE: 12-SEP-1997
; APPLICATION NUMBER: JP 27,148/1996
; FILING DATE: 19-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KORNBAU, Anne M.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TOKUNAGA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-077-098A-6

Query Match      7.5%; Score 174; DB 4; Length 2042;
Best Local Similarity 23.2%; Pred. No. 8e-05;
Matches 114; Conservative 64; Mismatches 189; Indels 124; Gaps 22;

```


Best Local Similarity 32.0%; Pred. No. 3.3e-05;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

Qy 18 QSGD---NGLGHNANSALGQOPIDRTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73
Db 134 QPGNDKNGVGGANGAKGAGGCGGLAEALQETIEQILQLGGGAGAGGAGGAGGAGGA 193

Qy 74 DQTTGVNAGGLNGRKTAGT--TPQSDSQNMLSEMGNGLDQAITPDG--QSGGQIGDN 129
Db 194 DGGSGAGGAGGANGADGGNGVNGNQANGPQAGDVNGANGAD-----DGSDDQGGTGV 248

Qy 130 PLLKMKILKLARMD-----GQSDQFGQPGCTGNNSSASSTSSGSP--FNDLSGCK 179
Db 249 QKLMKILNALVQMQOGLGGNQAGGSKGAGNAPASGANPCANQPGSADDSGQ 306

RESULT 13
US-09-030-270A-7
; Sequence 7, Application US/09030270A
; Patent No. 5977060
; GENERAL INFORMATION:
; APPLICANT: Zitter, Thomas A.
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: INSECT CONTROL WITH A
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,270A
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,226
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-030-270A-7

Query Match 7.2%; Score 165.5; DB 2; Length 344;
Best Local Similarity 32.0%; Pred. No. 3.3e-05;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

Qy 18 QSGD---NGLGHNANSALGQOPIDRTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73
Db 134 QPGNDKNGVGGANGAKGAGGCGGLAEALQETIEQILQLGGGAGGAGGAGGAGGAGGA 193

Qy 74 DQTTGVNAGGLNGRKTAGT--TPQSDSQNMLSEMGNGLDQAITPDG--QSGGQIGDN 129
Db 194 DGGSGAGGAGGANGADGGNGVNGNQANGPQAGDVNGANGAD-----DGSDDQGGTGV 248

Qy 130 PLLKMKILKLARMD-----GQSDQFGQPGCTGNNSSASSTSSGSP--FNDLSGCK 179
Db 249 QKLMKILNALVQMQOGLGGNQAGGSKGAGNAPASGANPCANQPGSADDSGQ 306

RESULT 14
US-08-984-207-7
; Sequence 7, Application US/08984207
; Patent No. 6235974
; GENERAL INFORMATION:
; APPLICANT: Qiu, Dewen
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Bear, Steven V.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
; TITLE OF INVENTION: RESISTANCE IN PLANTS BY SEED TREATMENT
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,207
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,230
; FILING DATE: 05-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1201
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-984-207-7

Query Match 7.2%; Score 165.5; DB 3; Length 344;
Best Local Similarity 32.0%; Pred. No. 3.3e-05;
Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;

Qy 18 QSGD---NGLGHNANSALGQOPIDRTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73
Db 134 QPGNDKNGVGGANGAKGAGGCGGLAEALQETIEQILQLGGGAGGAGGAGGAGGAGGA 193

Qy 74 DQTTGVNAGGLNGRKTAGT--TPQSDSQNMLSEMGNGLDQAITPDG--QSGGQIGDN 129
Db 194 DGGSGAGGAGGANGADGGNGVNGNQANGPQAGDVNGANGAD-----DGSDDQGGTGV 248

Qy 130 PLLKMKILKLARMD-----GQSDQFGQPGCTGNNSSASSTSSGSP--FNDLSGCK 179
Db 249 QKLMKILNALVQMQOGLGGNQAGGSKGAGNAPASGANPCANQPGSADDSGQ 306

RESULT 15
US-09-013-587-7
; Sequence 7, Application US/09013587
; Patent No. 6277814
; GENERAL INFORMATION:

QY 18 QSGD---NGLGHNANSALGQPIDRQTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73
 Db 134 QPGNDKNGVGGANGAKAGGQGLAEALQETIEQILAQIAGGAGAGAGGAGGAGGAGGA 193
 QY 74 DQTTGVNAGLNGRKTACT--TPQSDSONMLSEMNGNGLDQAITPDG--QGGGQIGDN 129
 Db 194 DGGGAGGAGGANGADGGVNGVNGQANGPQAGDVNGANGAD-----DGEDQGGTGV 248
 QY 130 PLLKMLKLIARMD-----GQSDQFGQPGTGNNSSASCTSSSGSP--FNDSLGGK 179
 Db 249 QKLMKILNALVQMQGGGLGGGNAQGGSKGAGNAPASGANPQAGSADQSSGQ 306

RESULT 18

PCT-US96-08819-7
 ; Sequence 7, Application PC/TUS9608819
 ; GENERAL INFORMATION:
 ; APPLICANT: Cornell Research Foundation, Inc.
 ; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
 ; TITLE OF INVENTION: RESISTANCE IN PLANTS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 ; STREET: Clinton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/08819
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/475,775
 ; FILING DATE: 07-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldman, Michael L.
 ; REGISTRATION NUMBER: 30,727
 ; REFERENCE/DOCKET NUMBER: 19603/10051
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (716) 263-1304
 ; TELEFAX: (716) 263-1600
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 344 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US96-08819-7

Query Match 7.2%; Score 165.5; DB 5; Length 344;
 Best Local Similarity 32.0%; Pred. No. 3.3e-05;
 Matches 57; Conservative 20; Mismatches 80; Indels 21; Gaps 7;
 QY 18 QSGD---NGLGHNANSALGQPIDRQTIEQMAQLLAEI-LKSLSPQSGNAATGAGN 73
 Db 134 QPGNDKNGVGGANGAKAGGQGLAEALQETIEQILAQIAGGAGAGAGGAGGAGGAGGA 193
 QY 74 DQTTGVNAGLNGRKTACT--TPQSDSONMLSEMNGNGLDQAITPDG--QGGGQIGDN 129
 Db 194 DGGGAGGAGGANGADGGVNGVNGQANGPQAGDVNGANGAD-----DGEDQGGTGV 248
 QY 130 PLLKMLKLIARMD-----GQSDQFGQPGTGNNSSASCTSSSGSP--FNDSLGGK 179
 Db 249 QKLMKILNALVQMQGGGLGGGNAQGGSKGAGNAPASGANPQAGSADQSSGQ 306

RESULT 19
 US-09-010-928B-4
 ; Sequence 4, Application US/09010928B
 ; Patent No. 5994099
 ; GENERAL INFORMATION:
 ; APPLICANT: Lewis, Randolph V
 ; APPLICANT: Hayashi, Cheryl Y
 ; TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
 ; TITLE OF INVENTION: CODING THEREFOR
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 ; STREET: 8110 GATEHOUSE RD. SUITE 500E
 ; CITY: FALLS CHURCH
 ; STATE: VIRGINIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 22042
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/010,928B
 ; FILING DATE: 22-JAN-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murphy Jr., Gerald M
 ; REGISTRATION NUMBER: 28977
 ; REFERENCE/DOCKET NUMBER: 1447-109P
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 907 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-010-928B-4

Query Match 6.9%; Score 160.5; DB 2; Length 907;
 Best Local Similarity 21.4%; Pred. No. 0.00031;
 Matches 85; Conservative 35; Mismatches 134; Indels 143; Gaps 14;
 QY 12 SSPGLPQSGDNGLGHNNANSALGQPIDRQTIEQMAQLLAEI-LKSLSPQSGNA--ATG 69
 Db 603 SGPGGYGGG--SGAGGTGPGGAGG-----ACGAGSGG 634
 QY 70 AGNDQTTGCVNAGLNGRKTACTTPQSD-----SQNMLSEMNGNGLDQ 114
 Db 635 AGSGGAGGSGGAGGSGGSGGTITTEDLDTIDGADGPTITSELTISGAGSGPGG 694
 QY 115 A----ITPDGQGGQIGDNPLLKMLKLIARIMMDGQSDQFGQPGTGNNSSASCTSSSGG- 169
 Db 695 AGPGGVGPGSGPGGVGPG-----VSGPGG-VGPGSGGPGGVGSGSGPGGV 740
 QY 170 -----SPFNDLSGGKAPSGNSPSGN-----YSPVSTFSP-----PST 201
 Db 741 GPGGVGPGSGSGGVGPGGVGPGSGGFGYGGSGEGYGPSTGYGSGGGYGPAGGPGY 800
 QY 202 PPSPTSPDPSPPTKAAGSTPVTDPHPVGSAGIGAGNSVAFVSAGANQTVLHDTIV 261
 Db 801 PGSPGGAYG-PGSPGGAYYPSRVPDMVNGIMSAMQSGGFNY-----841
 QY 262 KAGQVPDGGQFTAGSELGDCGQSENQKPLFILEDGASLKNVTMGDGDGADGIHLXDAAK 321
 Db 842 ---QMPGNMLSQYSSSS-----GTCNPNVNVLMAL 870
 QY 322 IDNLH-VTVNGDAITVKPNSAGKSHVITNSSFH 357
 Db 871 LAALHCLSNHGSSSPAPSPPTPAAMSAYSNSVGRMFAY 907

RESULT 20

Query Match	6.8%;	Score 157.5;	DB 4;	Length 2870;
Best Local Similarity	20.8%;	Pred. No. 0.0026;		
Matches 103;	Conservative 62;	Mismatches 201;	Indels 129;	Gaps 17;

QY	3	ILTLN----	NNTSSPFLPQSGDNLGNGHNANSALGOQPIDRQTIQWQAALLAELKSL	58
Db	1079	IFTMNVLPPTTTTETPTPTTSSDDAGGKTGTGATG-----	1115	
QY	59	LSPQSNAAATGAGGNDQTGTGNA-----GGLNGRKGT-----	AGTTPQSDQNMLSE	106
Db	1116	-----GTGTTSGGGATTTLTSGDAVTRTSGSGSGQSTGSCAGSGGTTASGSGSGSG	1170	
QY	107	MGNGLDQ-----AITPDQGG-----	GGQIGDNPLLKAMLKLTARMMWDGSDQFGQPGT	155
Db	1171	TGSDGNSGKTTALNGDGTGSGTATTPGSHLGD-----	1203	
QY	156	GNNASGCTSSGSGSPFN-DLSGGKAPGNS-PSGNYSPTSPPTSPPTSPPLDFPS	213	
Db	1204	GGTSGSGSDNSGSGVSTKSSGSDTSGSDSSGANGAFSAYTAQSTRTTKTRSSLATV	1263	
QY	214	SPTRAAAGGSTPVTDHPDPVGS-AGIGAGNSVATFSAGANOTVLHDTITVKGAQVDFKGQ	272	
Db	1264	SPISAAEQAIIDAQADVMQLAGIMDG-SASNNLSLNTSSLLNQISSLPAADLVEAQS	1322	
QY	273	TFTAGSELGCGGSEENOKPFILEDGASLKNVTMGDDGADGHLGYDAKIDNLHVTN--	330	
Db	1323	LLSNTLKIPGVGNMSSVDVLKTLQDNIATNSELADEMAKVI-----TKLANVNM TSAQS	1377	
QY	331	-----CEDAITYKPSAKGKHVEITNSSFEHAS-----	DKTL	363
Db	1378	LNSVLSLDDIALKGSTVYTLGVSSTSKSDGTVAVIFYAGSYTLVSPRCTLSIYGSTI	1437	
QY	364	QLMADTNLSY-----DNVKAQDFTVFRINGCGQGNWDLNLSHISAEDGKFSFVKDSEG	418	
Db	1438	YLTGDTRASVKQLDGTVTVA---DTMLAAAIQIGMFATNGRTVQVEQKIDDKESLVSQ	1494	
QY	419	LNVTNSDSDISLDGVEN	433	
Db	1495	-NIMATMSGVGDVQS	1508	

Search completed: January 25, 2005, 12:08:36
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:06:11 ; Search time 56.25 Seconds
(without alignments)
159.435 Million cell updates/sec

Title: us-09-879-248-6_COPY_116_140
Perfect score: 128
Sequence: 1 ITPDCGGGGIGDNPILLKAMKLIA 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_23Sep04:.*
1: Genesecp1980s:.*
2: Genesecp1990s:.*
3: Genesecp2000s:.*
4: Genesecp2001s:.*
5: Genesecp2002s:.*
6: Genesecp2003as:.*
7: Genesecp2003bs:.*
8: Genesecp2004s:.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	447	2	AAW96260 Hypersens
2	128	100.0	447	3	AAW71094 Erwinia a
3	128	100.0	447	3	AAW84855 A hyperse
4	128	100.0	447	5	AAO22548 Hypersens
5	128	100.0	447	5	AAE18296 Erwinia a
6	128	100.0	447	5	AAE16448 E. amylov
7	128	100.0	447	5	AAW09226 Erwinia a
8	57	44.5	206	2	AAW29770 Malassezi
9	54	42.2	837	3	AAE12381 N-termina
10	54	42.2	837	5	AAE14656 Murine ST
11	54	42.2	837	6	ABU10480 Mouse STA
12	52	40.6	135	2	AAW56974 Partial P
13	52	40.6	135	2	AAW11979 Mycobacte
14	52	40.6	207	2	AAW61476 A. fumiga
15	52	40.6	210	8	ADMS7310 Modular a
16	52	40.6	347	8	ADMS7298 Modular a
17	49	38.3	110	5	ADK34437 Novel hum
18	49	38.3	164	8	ADJ25641 Voiteage-g
19	49	38.3	436	8	ADN73559 Thale cre
20	49	38.3	613	6	ABR42419 Human pot
21	48.5	37.9	55	5	ABPO2086 Human ORF
22	48.5	37.9	193	8	ADN99445 Novel hum
23	48.5	37.9	365	4	ABG60270 Human ova
24	48.5	37.9	365	5	ABG61741 Novel ova
25	48.5	37.9	392	4	ADM19994 Protein e

26	48	37.5	43	7	ADN59761	Adn59761 Peptide-v
27	48	37.5	238	5	ABP73460	Abp73460 Candida a
28	48	37.5	505	4	ABB58292	Abb58292 Drosophil
29	47	36.7	72	7	ADF58708	Adf58708 Human pol
30	47	36.7	451	4	ABG10570	Abg10570 Novel hum
31	47	36.7	737	7	ADF83524	Adf83524 C. reinha
32	47	36.7	737	7	ADF83526	Adf83526 C. reinha
33	47	36.7	872	4	ABB69476	Abb69476 Drosophil
34	47	36.7	1343	4	AAU35444	Aau35444 Haemophil
35	47	36.7	1343	6	ABU30272	Abu30272 Protein e
36	47	36.7	1479	4	ABB70936	Abb70936 Drosophil
37	46.5	36.3	353	8	ADJ43241	Adj43241 Plant tra
38	46.5	36.3	353	8	ADO62808	Ado62808 Transcrip
39	46.5	36.3	504	4	ABB66600	Abb66600 Drosophil
40	46	35.9	402	3	AAO6450	Aao6450 Arabidops
41	46	35.9	402	6	ADA33693	Ada33693 Acinetoba
42	46	35.9	403	3	AAO43157	Aao43157 Arabidops
43	46	35.9	408	4	AAW79473	Aaw79473 Corynebac
44	46	35.9	419	7	ADB70132	Adb70132 C. neofor
45	46	35.9	434	3	AAO66449	Aao66449 Arabidops
46	46	35.9	435	3	AAO43156	Aao43156 Arabidops
47	46	35.9	441	4	AAO92180	Aao92180 C glutami
48	46	35.9	749	2	AAW70233	Aaw70233 P. falcip
49	46	35.9	749	2	AAW22479	Aaw22479 Plasmodiu
50	46	35.9	749	3	AAW77901	Aaw77901 P. falcip
51	46	35.9	828	4	ABB67048	Abb67048 Drosophil
52	46	35.9	1086	5	AAU76760	Aau76760 Plasmodiu
53	46	35.9	1098	7	ABO71440	Ab071440 Pseudomon
54	46	35.9	1150	4	ABB63215	Abb63215 Drosophil
55	45.5	35.5	72	5	ABP02813	Abp02813 Human ORF
56	45.5	35.5	197	4	AAU63166	Aau63166 Propionib
57	45.5	35.5	197	6	ABM59685	Abm59685 Propionib
58	45	35.2	23	7	ADN59774	Adn59774 Peptide-v
59	45	35.2	41	7	ADN59768	Adn59768 Peptide-v
60	45	35.2	46	7	ADN59780	Adn59780 Peptide-v
61	45	35.2	46	7	ADN59786	Adn59786 Peptide-v
62	45	35.2	113	5	ABP41357	Abp41357 Human ova
63	45	35.2	115	4	AAO13141	Aao13141 Human pol
64	45	35.2	133	7	ADF06199	Adf06199 Bacterial
65	45	35.2	253	3	AAO43442	Aao43442 Arabidops
66	45	35.2	253	3	AAO66682	Aao66682 Arabidops
67	45	35.2	257	3	AAO43441	Aao43441 Arabidops
68	45	35.2	257	3	AAO66681	Aao66681 Arabidops
69	45	35.2	69	2	AAW17969	Aaw17969 X. laevis
70	45	35.2	365	2	AAW53102	Aaw53102 Pig H tra
71	45	35.2	365	2	AAW30630	Aaw30630 Swine alp
72	45	35.2	365	2	AAW97356	Aaw97356 Swine alp
73	45	35.2	365	5	AAW79302	Aaw79302 Pig alpha
74	45	35.2	365	5	AAW47995	Aaw47995 Swine alp
75	45	35.2	403	4	ABB63509	Abb63509 Drosophil
76	45	35.2	515	2	AAW54236	Aaw54236 Human Lam
77	45	35.2	515	4	AAW90770	Aaw90770 Human she
78	45	35.2	532	6	ABJ19005	Abj19005 Pathogen
79	45	35.2	532	6	ABM72976	Abm72976 Staphyloc
80	45	35.2	581	6	ABP65400	Abp65400 Bifidobac
81	45	35.2	664	6	ABU04992	Abu04992 Human exp
82	45	35.2	664	7	ADJ63975	Adj63975 Human pro
83	45	35.2	664	7	ADJ70273	Adj70273 Human hea
84	45	35.2	664	8	ADQ13765	Ado13765 Human wil
85	45	35.2	665	7	ADJ63973	Adj63973 Rat Prote
86	45	35.2	689	7	ABB64493	Abb64493 Drosophil
87	45	35.2	789	6	ABU21326	Abu21326 Protein e
88	45	35.2	874	6	ABU21321	Abu21321 Protein e
89	45	35.2	1254	4	ABB62006	Abb62006 Pseudomon
90	45	35.2	1403	7	ABO76754	Ab076754 Pseudomon
91	45	35.2	1547	4	ABB58643	Abb58643 Drosophil
92	45	35.2	1780	4	ABB71483	Abb71483 Drosophil
93	44.5	34.8	202	4	ABG05876	Abg05876 Novel hum
94	44.5	34.8	928	8	ADA04052	Ada04052 Human IC-
95	44.5	34.8	1151	6	ADA34589	Ada34589 Acinetoba
96	44	34.4	97	4	AAO03289	Aao03289 Human pol
97	44	34.4	110	4	AAO63951	Aao63951 Human pro
98	44	34.4	124	3	AAW28294	Aaw28294 Arabidops

99 44 34.4 200 3 AAG28293 Aag28293 Arabidops
100 44 34.4 203 4 AAU64010 Aau64010 Propionib

ALIGNMENTS

RESULT 1

AAW96260
ID AAW96260 standard; protein; 447 AA.

XX
AC AAW96260;
XX

DT 14-JUN-1999 (first entry)
XX

DE Hypersensitive response eliciting protein (HrpW).
XX

KW Hypersensitive response elicitor protein; hypersensitive response; hrpW;
KW pathogen; infection; crop protection; disease resistance;
KW pest resistance; transgenic plant; colouration; maturation.
XX

OS Erwinia amylovora.
XX

PN WO9907208-A1.
XX

PD 18-FEB-1999.
XX

XX 27-JUL-1998; 98WO-US015547.
XX

PF 06-AUG-1997; 97US-0055108P.
XX

PR (CORR) CORNELL RES FOUND INC.
XX

PA Kim JF, Beer SV;
XX

PI WPI; 1999-167126/14.
XX

DR N-PSDB; AAX09007.
XX

XX New Erwinia amylovora hypersensitive response eliciting gene and protein
PT - useful for providing transgenic plants and seeds with enhanced growth,
PT and insect and disease resistance.
XX

PS Claim 1; Page 50-51; 54pp; English.
XX

XX The hypersensitive response eliciting protein (hrp) or polypeptide is
CC produced as part of an active defense by plants against incompatible
CC pathogen infections. The hypersensitive response is a rapid localised
CC necrosis. The hrp protein and gene when used in nucleotide constructs are
CC useful for providing disease resistance to plants, insect control to
CC plants, and enhancing plant growth (enhancing fruit size and earlier
CC colouration and maturation), by direct application of the protein to
CC plants, or by producing transgenic plants or seeds using the hrp gene
XX

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 2; Length 447;

Best Local Similarity 100.0%; Pred. No. 5.9e-11;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLLKAMLKLI A 25

DB 116 ITPDGGGGGQIGDNPLLKAMLKLI A 140

RESULT 2

AAV71094

ID AAV71094 standard; protein; 447 AA.

XX
AC AAV71094;
XX

DT 08-SEP-2000 (first entry)
XX

DE Erwinia amylovora hypersensitive response elicitor #2.

XX Hypersensitive response elicitor; environmental stress resistance; plant.
KW Erwinia amylovora.
OS

PN WO200028055-A2.
XX

PD 18-MAY-2000.
XX

PF 04-NOV-1999; 99WO-US026039.
XX

PR 05-NOV-1998; 98US-0107243P.
XX

PA (EDEN-) EDEN BIOSCIENCE CORP.
XX

PI Wei Z, Schading RL;
XX

XX WPI; 2000-376566/32.
XX

DR N-PSDB; AAD00669.
XX

XX Application of a hypersensitive response elicitor protein to plants to
PT impart stress resistance.
XX

PS Disclosure; Page 10-12; 84pp; English.
XX

XX The patent discloses a method to impart stress resistance to plants by
CC applying a hypersensitive response elicitor in a non-infectious form to a
CC plant or seed. The present sequence is a hypersensitive response elicitor
CC protein from Erwinia amylovora. The protein is heat stable, protease
CC sensitive and suppressed by inhibitors of plant metabolism. It is used to
CC impart stress resistance to plants
XX

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 3; Length 447;

Best Local Similarity 100.0%; Pred. No. 5.9e-11;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPLLKAMLKLI A 25

DB 116 ITPDGGGGGQIGDNPLLKAMLKLI A 140

RESULT 3

AAV84855

ID AAV84855 standard; protein; 447 AA.

XX
AC AAV84855;
XX

DT 08-AUG-2000 (first entry)
XX

DE A hypersensitive response elicitor protein.
XX

KW Hypersensitive response; insect control; disease resistance;
KW hypersensitive response elicitor; plant growth; vegetable; crop;
KW ornamental plant.
XX

OS Erwinia amylovora.
XX

PN WO200020452-A2.
XX

PD 13-APR-2000.
XX

PF 05-OCT-1999; 99WO-US023181.
XX

PR 05-OCT-1998; 98US-0103050P.
XX

PA (EDEN-) EDEN BIOSCIENCE CORP.
XX

PI Wei Z, Fan H, Niggemeyer JL;
XX

XX WPI; 2000-303745/26.
XX

DR N-PSDB; AAA14939.

XX Hypersensitive response elicitor polypeptides useful for imparting
PT enhanced growth, disease resistance and insect resistance to plants,
PT especially vegetables and ornamental flowers.
XX
XX PS Disclosure; Page 12-13; 100pp; English.
XX
XX The present sequence represents a hypersensitive response elicitor
CC polypeptide. The specification describes hypersensitive response elicitor
CC polypeptide fragments, which do not elicit a hypersensitive response.
CC Instead, the proteins impart disease resistance to plants, enhance plant
CC growth, and/or control insects. The polypeptide fragments may be used to
CC these properties to plants. The plants which may be treated in this way
CC include vegetables, crops and ornamental plants such as alfalfa, rice,
CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet
CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussels sprout,
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,
CC carnation or zinnia
XX
XX Sequence 447 AA;
Query Match 100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ITPDGGGGGQIGDNPLLKAMKLIIA 25
DB 116 ITPDGGGGGQIGDNPLLKAMKLIIA 140
RESULT 4
AAO22548
ID AAO22548 standard; protein; 447 AA.
XX
XX AAO22548;
XX
XX 28-OCT-2002 (first entry)
XX
XX Hypersensitive response elicitor protein #2 from Erwinia amylovora.
XX
XX Desiccation inhibitor; longevity enhancer; desiccation inhibitor;
KW cutting; ornamental plant; hypersensitive response elicitor protein;
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;
KW flower bloom; flower.
XX
XX Erwinia amylovora.
OS
XX WO200237960-A2.
PN
XX 16-MAY-2002.
PD
XX
XX 06-NOV-2001; 2001WO-US043715.
PF
XX
XX 13-NOV-2000; 2000US-0248169P.
PR
XX
XX (EDEN-) EDEN BIOSCIENCE CORP.
PA
XX
XX Wei Z, Leon E, Oviedo A;
PI
XX WPI: 2002-575194/61.
DR
XX N-PSDB; AAL41133.
DR
XX Inhibiting desiccation of cuttings from ornamental plants, by treating
PT ornamental plants with hypersensitive response elicitor protein, or
PT expressing heterologous hypersensitive response elicitor protein in
PT plants.
XX
XX Disclosure; Page 12-13; 69pp; English.
XX

CC The invention relates to a method for inhibiting desiccation of cuttings
CC from ornamental plants. The method involves treating the cuttings with a
CC hypersensitive response elicitor protein or polypeptide, or providing a
CC transgenic ornamental plant or plant seed transformed with a DNA molecule
CC encoding the hypersensitive response elicitor polypeptide, and growing
CC the ornamental plant or transgenic ornamental plant produced from the
CC transgenic ornamental plant seed. The hypersensitive response elicitor
CC protein or polypeptide is useful for inhibiting desiccation of cuttings
CC from ornamental plants, for harvesting cuttings from ornamental plants,
CC for promoting early flowering of ornamental plants, and enhancing the
CC longevity of flower blooms on ornamental plant cuttings. This sequence
CC represents a hypersensitive response elicitor protein of the invention
XX
XX Sequence 447 AA;
Query Match 100.0%; Score 128; DB 5; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ITPDGGGGGQIGDNPLLKAMKLIIA 25
DB 116 ITPDGGGGGQIGDNPLLKAMKLIIA 140
RESULT 5
AAE18296
ID AAE18296 standard; protein; 447 AA.
XX
XX AAE18296;
XX
XX 07-MAY-2002 (first entry)
XX
XX Erwinia amylovora hypersensitive response elicitor (HRE) #2.
DE
XX
XX Hypersensitive response elicitor; HRE; transgenic plant; plant growth;
KW stress tolerance; disease tolerance; modified flower colour;
KW insect resistance; herbicide resistance; male sterility.
XX
XX Erwinia amylovora.
OS
XX WO200195724-A2.
PN
XX 20-DEC-2001.
PD
XX
XX 13-JUN-2001; 2001WO-US018955.
PF
XX
XX 15-JUN-2000; 2000US-0211585P.
PR
XX
XX (EDEN-) EDEN BIOSCIENCE CORP.
PA
XX
XX Wei Z, Derocher J;
PI
XX WPI: 2002-130707/17.
DR
XX N-PSDB; AAD29125.
DR
XX
XX Improving effectiveness of transgenic plants by topical application of a
PT hypersensitive response elicitor protein to the transgenic plant or by
PT incorporating into the plant a transgene encoding the protein.
XX
XX Disclosure; Page 13-14; 86pp; English.
XX
XX The invention relates to methods of improving the effectiveness of
CC transgenic plants which involves either topical application of a
CC hypersensitive response elicitor (HRE) protein to the transgenic plant or
CC incorporating into the transgenic plant a transgene encoding HRE. HRE
CC sequence is used for improving the effectiveness of transgenic plants by
CC maximising the benefit of transgenic traits associated with a deleterious
CC effect on growth, stress tolerance, disease or insect resistance,
CC enhanced growth, herbicide resistance, male sterility, modified flower
CC colour and biochemically modified plant product in the transgenic plants
CC or overcoming the deleterious effects. The present sequence is Erwinia
CC amylovora HRE protein
XX

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAMLKLI A 25
 |||||
 Db 116 ITPDGGGGQIGDNPPLKAMLKLI A 140
 |||||

RESULT 6
 AAEL16448
 ID AAEL16448 standard; protein; 447 AA.
 AC AAEL16448;
 DT 09-APR-2002 (first entry)
 DE E. amylovora hypersensitive response elicitor protein #2.
 KW Hypersensitive response elicitor protein; plant growth; fruit coloration;
 KW disease resistance; stress resistance; phytotoxin; insect infection;
 KW plant maturation.
 OS Erwinia amylovora.
 FH Key
 FT Location/Qualifiers
 FT Domain
 FT 5..64
 FT /label= Hypersensitive_response_eliciting_domain_1
 FT Region
 FT 5..45
 FT /label= Acidic_unit
 FT Region
 FT 45..64
 FT /label= Alpha_helix
 FT Domain
 FT 103..146
 FT /label= Hypersensitive_response_eliciting_domain_2
 FT Region
 FT 103..131
 FT /label= Acidic_unit
 FT Region
 FT 131..146
 FT /label= Alpha_helix
 XX WO200198501-A2.
 XX 27-DEC-2001.
 XX 12-JUN-2001; 2001WO-US018620.
 XX 16-JUN-2000; 2000US-0212211P.
 XX (EDEN-) EDEN BIOSCIENCE CORP.
 XX Fan H, Wei Z;
 XX WPI; 2002-122282/16.
 XX N-PSDB; AAD27016.
 XX New hypersensitive response elicitor proteins comprising spaced apart
 PT domains having an acidic portion linked to an alpha-helix, useful for
 PT imparting disease or stress resistance, controlling insects or enhancing
 PT plant growth.
 XX Disclosure; Page 13-14; 99pp; English.
 PS The patent discloses hypersensitive response elicitor proteins and
 XX nucleotides encoding such proteins. Hypersensitive response elicitor
 CC proteins comprise an isolated pair or more of spaced apart domains, each
 CC comprising an acidic portion linked to an alpha-helix and capable of
 CC eliciting a hypersensitive response in plants. Sequences of the invention
 CC are used to impart disease resistance to plants. Sequences of the invention
 CC to control insects and/or to impart stress resistance to plants which
 CC includes resistance to environmental stresses such as climate, air
 CC pollution, chemical and nutritional stress. The method of imparting
 CC disease resistance has the potential for treating previously untreatable

CC diseases, treating diseases systemically and avoiding the use of
 CC infectious agents or environmentally harmful materials. Hyper- sensitive
 CC response elicitor sequences are used to enhance plant growth which
 CC encompasses greater yield, increased in quantity of seeds produced,
 CC percentage of seeds germinated, plant size and biomass, bigger fruits,
 CC earlier fruit coloration and plant maturation. They are also used for
 CC insect control which encompasses preventing direct insect damage to plant
 CC by feeding injury, interfering with insect larval feeding on the plants,
 CC preventing insects from colonising host plants and releasing phytotoxins.
 CC Sequences of the invention also prevent subsequent disease damage to
 CC plants resulting from insect infection. The present sequence is Erwinia
 CC amylovora hypersensitive response elicitor protein

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPPLKAMLKLI A 25
 |||||
 Db 116 ITPDGGGGQIGDNPPLKAMLKLI A 140
 |||||

RESULT 7
 ABB09226
 ID ABB09226 standard; protein; 447 AA.
 AC ABB09226;
 XX 08-JUL-2002 (first entry)
 DT Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.
 DE Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;
 DE postharvest disease.
 KW Erwinia amylovora.
 OS WO200180639-A2.
 XX 01-NOV-2001.
 XX 17-APR-2001; 2001WO-US012468.
 XX 19-APR-2000; 2000US-0198359P.
 XX (EDEN-) EDEN BIOSCIENCE CORP.
 XX Wei Z, Qiu D, Remick D;
 XX WPI; 2002-041357/05.
 XX N-PSDB; ABL51711.
 XX Inhibiting post harvest disease (caused by Penicillium, Botrytis,
 PT Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a
 PT fruits or vegetables, using hypersensitive response elicitor proteins or
 PT nucleic acids.
 XX Disclosure; Page 11-12; 72pp; English.
 PS The present invention describes methods for inhibiting post harvest
 XX disease or desiccation and enhancing the longevity in a fruits or
 CC vegetables, using hypersensitive response elicitor proteins or
 CC polypeptides or nucleic acids (I) derived from pathogens (e.g. Erwinia
 CC amylovora, E. stewartii, E. chrysanthemi, E. carotovora, Xanthomonas,
 CC Pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter).
 CC (I) has bactericidal activity, and can be used in gene therapy. The
 CC method can be used for inhibiting post harvest disease (caused by
 CC Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and
 CC enhancing the longevity in a fruits or vegetables. The method enables
 CC growers, warehouse packers, shippers and suppliers to process, handle and
 CC store fruit and vegetables with reduced losses caused by post harvest

CC disease and desiccation, therefore reducing costs to the consumer and
 CC improving quality. The present sequence represents a hypersensitive
 CC response elicitor protein given in the exemplification of the present
 CC invention
 CC
 SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGNPLKAMKLIA 25
 |||||
 Db 116 ITPDGGGGGQIGNPLKAMKLIA 140

RESULT 8
 AAW29770
 ID AAW29770 standard; protein; 206 AA.

AC AAW29770;
 DT 27-AUG-2003 (revised)
 DT 23-FEB-1998 (first entry)
 XX Malassezia fungus MF-3 antigenic protein.

XX Malassezia; fungus; antigenic; human; IgE; immunoglobulin E; antibody;
 KW allergy; antigen.
 XX Malassezia.

XX WO9721817-A1.

XX 19-JUN-1997.

XX 10-DEC-1996; 96WO-JP003602.

XX 12-DEC-1995; 95JP-00346627.

XX 05-SEP-1996; 96JP-00257612.

XX 05-SEP-1996; 96JP-00257613.

XX (TAKI) TAKARA SHUZO CO LTD.

XX Takesako K, Okado T, Yagihara T, Kuroda M, Onishi Y, Kato I;

XX Akiyama K, Yasueda H, Yamaguchi H;

XX WPI; 1997-332788/30.

XX N-PSDB; AAT85876.

XX Claim 36; Page 83-84; 162pp; Japanese.

XX The present sequence represents a specifically claimed antigenic protein
 CC isolated from the fungus Malassezia. The antigenic protein can bind to
 CC IgE antibodies present in patients with allergic conditions. Antigenic
 CC proteins, peptides and nucleic acids from the fungus Malassezia can be
 CC used in the diagnosis, treatment and prevention of allergic conditions
 CC to Malassezia organisms (such as M.furfur, M. sympodialis and
 CC M.pachydermatitis). (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 206 AA;

Query Match 44.5%; Score 57; DB 2; Length 206;
 Best Local Similarity 45.0%; Pred. No. 2.1;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGNPLKAM 20

Db 86 LAPOSEGGGQINDGPLKQAI 105

RESULT 9
 AAB12381
 ID AAB12381 standard; peptide; 837 AA.
 XX AAB12381;
 AC AAB12381;
 XX 08-NOV-2000 (first entry)
 DT N-terminal domain of murine STAT-6 protein.

DE STAT; signal transducer and activator of transcription; crystal;
 KW drug design; murine.
 XX Mus sp.

XX Key Location/Qualifiers
 XX 4..9
 FT Region /label= Alpha helix 1
 FT 11..18
 FT Region /label= Alpha helix 2
 FT 24..29
 FT Region /label= Alpha helix 3
 FT 31..36
 FT Region /label= Alpha helix 4
 FT 39..43
 FT Region /label= Alpha helix 5
 FT 50..73
 FT Region /label= Alpha helix 6
 FT 75..90
 FT Region /label= Alpha helix 7
 FT 93..113
 FT Region /label= Alpha helix 8

XX US6087478-A.

XX 11-JUL-2000.

XX 23-JAN-1998; 98US-00012710.

XX 23-JAN-1998; 98US-00012710.

XX (UYRQ) UNIV ROCKEFELLER.

XX Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J;

XX WPI; 2000-505108/45.

XX New crystals of an N-terminal fragment of a signal transducer and
 PT activator of transcription that effectively diffracts x-rays, useful for
 PT drug screening and development.

XX Disclosure; Fig 1; 42pp; English.

XX The present invention relates to a crystal of an N-terminal fragment of a
 CC signal transducer and activator of transcription (STAT) protein. The
 CC crystal effectively diffracts x-rays, allowing the determination of the
 CC atomic coordinates of the N-terminal domain to a resolution of greater
 CC than 5.0 Angstroms. The present sequence is the N-terminal domain of the
 CC murine STAT 6 protein. The N-terminal domain enables STAT dimers to
 CC interact and bind DNA cooperatively, a mechanism important for gene
 CC activation. The crystals are useful in drug screening and development by
 CC selecting a potential drug by performing rational drug design with the 3-
 CC dimensional structure determined for the crystal

XX Sequence 837 AA;

Query Match 42.2%; Score 54; DB 3; Length 837;
 Best Local Similarity 64.3%; Pred. No. 30;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 GQGGGQIGNPLK 18

Qy 5 GGGGQIGDNPLK 18
 | : | | : | | |
Db 802 GEGGSLGSQPPLK 815

Query Match 42.2%; Score 54; DB 6; Length 937;
 Best Local Similarity 64.3%; Pred. No. 30;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGGGGIGDNPPLK 18
 | : | | | : | | | |
 Db 802 GEGGGLGSQLK 815

RESULT 12

AAR56974
 ID AAR56974 standard; protein; 135 AA.

XX AC AAR56974;

DT 28-FEB-1995 (first entry)

DE Partial PMGA 1.3 protein of Mycoplasma gallisepticum.

XX PMGA; adhesin gene complex; hemagglutinin; conserved sequences; primers;
 KW probes; amplification; polymerase chain reaction; specific; detection;
 KW PCR; 1.2; 1.3.

XX Mycoplasma gallisepticum.

XX Key Location/Qualifiers
 FT Peptide 1. .25
 FT /label= putative signal peptide

FT Protein 26.135

FT /label= mature PMGA1.2 protein

XX AU9350593-A.

PN

XX

PD 26-MAY-1994.

XX 10-NOV-1993; 93AU-00050593.

PR 10-NOV-1992; 92AU-00005744.

XX (UYME) UNIV MELBOURNE.

PA Browning GF, Markham PF, Whithear KG, Walker ID, Glew MD;

XX WPI; 1994-209061/26.

DR N-PSDB; AAQ86668.

XX Recombinant DNA constructs for Mycoplasma gallisepticum - for diagnosis,

PT treatment and prophylaxis of poultry respiratory disorders.

XX Example 1; Fig 3; 51pp; English.

XX AAR56974 shows the partial amino acid sequence of PMGA1.3 derived from

CC the partial DNA sequence of a 10 kb insert from a Mycoplasma

CC gallisepticum genomic DNA library, detected by probes based on the T3 and

CC C7 peptides. The proteins derived show high homology to PMGA. This

CC sequence is the best approximation of the sequence given in the

CC specification due to poor print quality. Mycoplasma gallisepticum

CC infection in poultry, humans and other animals is of economic importance

CC to many industries and it is desirable to produce effective vaccines and

CC probes for its detection. The sequences and probes and vaccine vectors of

CC the invention can be used for the diagnosis and treatment of Mycoplasma

CC gallisepticum infection, and for prophylaxis

XX

SQ Sequence 135 AA;

Query Match 40.6%; Score 52; DB 2; Length 135;

Best Local Similarity 64.3%; Pred. No. 7.8;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGQGGGIGDNP 15

| : | | | : | | | |

Db 57 TNPNGGGGMMGDNP 70

RESULT 13

AAW11979

ID AAW11979 standard; protein; 135 AA.

XX AC AAW11979;

DT 16-OCT-2003 (revised)

DT 23-APR-1997 (first entry)

XX Mycobacterium gallisepticum PMGA1.3 adhesin partial sequence.

XX Adhesin; PMGA; mycoplasma; diagnosis; vaccine; vector;

KW respiratory disease; poultry; haemagglutinin.

XX Mycoplasma gallisepticum; strain S6.

XX Key Location/Qualifiers
 FT Peptide 1. .25
 FT /label= sig_peptide

FT /note= "the signal peptide shows homology to the PMGA1.2
 signal peptide"

PN CA2135330-A.

XX 11-MAY-1995.

XX 08-NOV-1994; 94CA-02135330.

XX 10-NOV-1993; 93AU-00050593.

PR 20-APR-1994; 94US-00230312.

XX (BROW/) BROWNING G F.

XX Browning GF, Markham PF, Whithear KG, Walker ID, Glew MD;

XX WPI; 1995-241027/32.

DR N-PSDB; AAT51531.

XX New promoter region from a Mycoplasma gallisepticum adhesin gene - useful

PT when coupled to foreign antigen gene, for prodn. of multivalent live

PT vaccines, also new probes for detecting Mycoplasma and manipulating its

PT genome.

PT Disclosure; Fig 3; 81pp; English.

XX Adhesin PMGA1.2 (AAW11978) and adhesin PMGA1.3 fragment (AAW11979) are

CC products of gene sequences (see also AAT51531) isolated from Mycoplasma

CC gallisepticum. DNA constructs incorporating the promoter and/or signal

CC sequences of the PMGA genes can be used in the prodn. of multivalent live

CC vaccines. The signal peptide sequence is utilised where attachment of an

CC exogenous antigen gene to the mycoplasma cell membrane is required.

CC (Updated on 16-OCT-2003 to standardise OS field)

XX

SQ Sequence 135 AA;

Query Match 40.6%; Score 52; DB 2; Length 135;

Best Local Similarity 64.3%; Pred. No. 7.8;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPDGQGGGIGDNP 15

| : | | | : | | | |

Db 57 TNPNGGGGMMGDNP 70

RESULT 14

AAW61476

ID AAW61476 standard; protein; 207 AA.

XX AC AAW61476;

DT 19-OCT-1998 (first entry)

XX The present invention relates to a modular antigen-transporting molecule
 CC (MAT) comprising one each of a translocation module, a targeting module
 CC and an antigen module. MAT molecules and their coding sequences are used
 CC for imaging, and as pharmaceutical, vaccinating and diagnostic agents,
 CC for preventing, reducing and/or stimulating the immune response, and for
 CC treatment of autoimmune, allergic, rheumatism, organ rejection, infection
 CC (bacterial, viral or caused by eukaryotic pathogens) and/or malignant
 CC disease. The present sequence is a MAT molecule shown in the
 CC exemplification of the invention.
 XX
 SQ Sequence 347 AA;
 Query Match 40.6%; Score 52; DB 8; Length 347;
 Best Local Similarity 40.9%; Pred. No. 23;
 Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ITPDGGGGGQIGDNPVLKMLK 22
 Db 221 LAPEKSGGKIDQAPVLKAAIE 242
 RESULT 17
 ADK34437
 ID ADK34437 standard; protein; 110 AA.
 XX
 AC ADK34437;
 DT 06-MAY-2004 (first entry)
 XX
 DE Novel human polypeptide SeqID6519.
 XX
 KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;
 KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
 KW fungus; parasite; human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..110
 FT /label= OTHER
 FT /note= "OTHER= All Xaa's in this sequence are unknown
 FT amino acids or the site of a stop codon within the DNA
 FT sequence"
 XX
 PN WO200216439-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 05-MAR-2001; 2001WO-US004941.
 XX
 PR 07-MAR-2000; 2000US-00519705.
 PR 19-MAY-2000; 2000US-00574454.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2002-280918/32.
 XX
 PT Isolated polynucleotide encoding bone marrow derived polypeptides useful
 PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
 PT disease, and inflammatory bowel disease.
 XX
 PS Claim 20; SEQ ID NO 6519; 504pp; English.
 XX
 CC This invention relates to a novel isolated polynucleotide comprising a
 CC nucleotide sequence selected from one of 1680 sequences, a mature protein
 CC coding portion of them, an active domain of them and their complementary
 CC sequences. The invention may be useful for the production of compounds

CC with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
 CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory,
 CC antibacterial, antiviral, antifungal or antiparasitic activity. In
 CC addition, the disclosed sequences may be useful for gene therapy. The
 CC polypeptides or their antibodies are useful for treating many diseases
 CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
 CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
 CC viruses, fungi or parasites. The present sequence is that of a human
 CC polypeptide of the invention.
 XX
 SQ Sequence 110 AA;
 Query Match 38.3%; Score 49; DB 5; Length 110;
 Best Local Similarity 61.1%; Pred. No. 18;
 Matches 11; Conservative 1; Mismatches 2; Indels 4; Gaps 1;
 QY 3 PDGGGGGQIG----DNPL 16
 Db 10 PNGGGGGQIGAHGAPNPL 27
 RESULT 18
 ADJ25641
 ID ADJ25641 standard; protein; 164 AA.
 XX
 AC ADJ25641;
 DT 20-MAY-2004 (first entry)
 XX
 DE Voltage-gated potassium channel amino acid sequence hKv1.5.
 XX
 KW potassium channel; voltage-gated potassium channel; mutant; yeast;
 KW detection; neuronal integration; cardiac pacemaking; muscle contraction;
 KW hormone secretion; cell volume regulation; lymphocyte differentiation;
 KW cell proliferation.
 XX
 OS Homo sapiens.
 XX
 PN WO2004008103-A2.
 XX
 PD 22-JAN-2004.
 XX
 PF 14-JUL-2003; 2003WO-US022165.
 XX
 PR 12-JUL-2002; 2002US-0395272P.
 XX
 PA (CARD-) CARDIOME INC.
 XX
 PI Fedida D, Steele D;
 XX
 DR WPI; 2004-156549/15.
 XX
 PT New potassium channel comprising mutations producing an open voltage-
 PT gated channel, useful in information transfer and synaptic functions in
 PT neurons and participates in neuronal integration, cardiac pacemaking or
 PT muscle contraction.
 XX
 PS Disclosure; Fig 4; 36pp; English.
 XX
 CC The present invention describes a potassium channel comprising a voltage-
 CC gated potassium channel which when expressed in a mutant yeast deficient
 CC in potassium uptake allows the mutant yeast to grow in the presence of
 CC media with very low potassium concentration, where the voltage-gated
 CC potassium channel comprises one or more mutations which produces a
 CC constitutively open voltage-gated potassium channel. Also described: (1)
 CC a yeast cell comprising a deficiency in potassium uptake and a
 CC constitutively open voltage-gated potassium channel which allows the
 CC yeast cell to grow in the presence of media with very low potassium; (2)
 CC a recombinant nucleic acid molecule comprising a nucleic acid sequence
 CC encoding the potassium channel and a promoter sequence operably linked to
 CC the nucleic acid molecule; and (3) a method of identifying compounds that
 CC modulates activity of voltage-gated ion channels. The potassium channel
 CC can be used as a detection system for detecting inhibitors and/or

CC activators of normally voltage-gated ion channels. The channels play a
 CC role in information transfer and synaptic functions in neurons and
 CC participate in neuronal integration, cardiac pacemaking, muscle
 CC contraction, hormone secretion, cell volume regulation, lymphocyte
 CC differentiation and cell proliferation. The present invention describes a
 CC voltage-gated potassium channel, which is given in the exemplification of
 CC the present invention.

XX Sequence 164 AA;

Query Match 38.3%; Score 49; DB 8; Length 164;
 Best Local Similarity 45.5%; Pred. No. 28;
 Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPLKAMKLII 24
 ||||| : : : : :
 Db 15 PGGGGGGGQGGQAMSLAIRVI 36

RESULT 19

ADN73559
 ID ADN73559 standard; protein; 436 AA.

XX AC ADN73559;

XX DT 15-JUL-2004 (first entry)

XX DE Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 1454.
 XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
 XX KW animal feed product; thale cress; cell wall biosynthesis;
 XX KW nitrogen metabolism; carbon metabolism.

XX OS Arabidopsis thaliana.

XX PN WO2004035798-A2.

XX PD 29-APR-2004.

XX PF 20-OCT-2003; 2003WO-EP011658.

XX PR 18-OCT-2002; 2002EP-00079408.

XX PA (CROP-) CROPDISEIGN NV.

XX PI Inze D, De Veylder L, Vlieghe K;

XX DR WPI; 2004-348466/32.

XX DR N-PSDB; ADN73558.

XX PT Altering plant characteristics, useful for producing plants for enzyme or
 XX PT pharmaceutical production comprising modifying in a plant, expression of
 XX PT one or more nucleic acids and/or modifying level or activity of one or
 XX PT more proteins.

XX PS Claim 1; SEQ ID NO 1454; 134pp; English.

XX CC This invention relates to a novel method for altering one or more plant
 XX CC characteristics. Specifically, it refers to identifying genes that are up
 XX CC or down-regulated in transgenic plants overexpressing the heterodimeric
 XX CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
 XX CC alter plant characteristics accordingly. The present invention describes
 XX CC generating transgenic plants for the production of growth regulators,
 XX CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
 XX CC the altered plant characteristics are selected from increased yield or
 XX CC biomass, enhanced survival capacity, stress tolerance, plant architecture
 XX CC or physiology, altered endoreplication, biochemistry, signal
 XX CC transduction, storage lipid mobilisation and/or altered photosynthesis,
 XX CC each relative to the corresponding wild type plants. Accordingly, these
 XX CC sequences can also be useful as positive or negative selectable markers
 XX CC during transformation of cells or tissues. The identified genes play a
 XX CC role in a variety of biological processes such as DNA replication, cell
 XX CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as

CC transcription factors. This polypeptide sequence is thale cress protein
 CC expressed by a gene repressed 1.3 fold or more in plants overexpressing
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the
 CC invention.

XX Sequence 436 AA;

Query Match 38.3%; Score 49; DB 8; Length 436;
 Best Local Similarity 52.6%; Pred. No. 84;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPLKAMKL 21
 ||||| : : : : :
 Db 22 PDKPGGGQINSNSVLVALL 40

RESULT 20

ABR42419
 ID ABR42419 standard; protein; 613 AA.

XX AC ABR42419;

XX DT 11-AUG-2003 (first entry)

XX DE Human potassium channel Kv1.5.

XX KW Human; potassium channel; Kv1.5; neuronal disease; multiple sclerosis;
 XX KW neuroprotective.

XX OS Homo sapiens.

XX PN WO2003035690-A2.

XX PD 01-MAY-2003.

XX PF 25-OCT-2002; 2002WO-GB004821.

XX PR 25-OCT-2001; 2001GB-00025636.

XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX PI Dolly JO, Akhtar S, Shamotienko O;

XX DR WPI; 2003-421404/39.

XX PT Preparing a K+ channel fusion protein comprising alpha subunits and beta
 XX PT subunits, useful in the manufacture of a medicament for treating multiple
 XX PT sclerosis.

XX PS Disclosure; Fig 4; 93pp; English.

XX CC The present sequence is the protein sequence of human potassium channel
 XX CC Kv1.5. The invention provides oligomeric potassium channel fusion
 XX CC proteins, preferably comprising Kv1.1 and/or Kv1.2 subunits, and methods
 XX CC of producing them in mammalian cells. The fusion proteins can be used to
 XX CC identify compounds that modulate potassium channel activity. Such
 XX CC modulator compounds are useful for the treatment of multiple sclerosis

XX Sequence 613 AA;

Query Match 38.3%; Score 49; DB 6; Length 613;
 Best Local Similarity 45.5%; Pred. No. 1.2e+02;
 Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPLKAMKLII 24
 ||||| : : : : :
 Db 381 PGGGGGGGQGGQAMSLAIRVI 402

Search completed: January 25, 2005, 12:12:07
 Job time : 65.25 secs


```

RESULT 2
US-09-431-614-6
; Sequence 6, Application US/09431614
; Patent No. 6624139
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/41 (EBC-003)
; CURRENT APPLICATION NUMBER: US/09/431,614
; CURRENT FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/107,243
; EARLIER FILING DATE: 1998-11-05

```

```
Query Match      44.5%; Score 57; DB 4; Length 206;
Best Local Similarity 45.0%; Pred. No. 0.6;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy   1  ITPDGGGGGIGINDPLIKAM 20
     :|::|||::||:|
Db    86 LAPOSEGGGLNDGPLKQAI 105
```



```
RESULT 7
US-09-248-796A-18831
; Sequence 18831, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18831
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18831

Query Match      37.5%; Score 48; DB 4; Length 252;
Best Local Similarity 43.5%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 9; Indels 9; Gaps 0;

QY      2 TPDGQGGGQIGDNPILKAMKL 24
Db      64 TTDGNGGGGIGIVPTLQIVATV 86

RESULT 8
US-09-270-767-46260
; Sequence 46260, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46260
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46260

Query Match      37.5%; Score 48; DB 4; Length 256;
Best Local Similarity 47.6%; Pred. No. 18;
Matches 10; Conservative 2; Mismatches 9; Indels 9; Gaps 0;

QY      4 DGGGGGQIGDNPILKAMKL 24
Db      40 DGDGGGGGGGVTLKAKMSLL 60

RESULT 9
US-08-576-626A-50
; Sequence 50, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L.
; APPLICANT: Donadio, S.
; APPLICANT: Staver, M.J.
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
; TITLE OF INVENTION: BIOSYNTHESIS GENES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories

; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/576,626A
; FILING DATE: 21-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne Casuto
; REGISTRATION NUMBER: P-40,943
; REFERENCE/DOCKET NUMBER: 5857.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (847) 938-3137
; TELEFAX: (847) 938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5998194e
US-08-576-626A-50

Query Match      36.7%; Score 47; DB 2; Length 379;
Best Local Similarity 56.2%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 5; Indels 5; Gaps 0;

QY      1 ITPDQGGGQIGDNP 16
Db      189 ITSGGEGGAVITDNPV 204

RESULT 10
US-09-328-352-4980
; Sequence 4980, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4980
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4980

Query Match      35.9%; Score 46; DB 4; Length 402;
Best Local Similarity 39.1%; Pred. No. 59;
Matches 9; Conservative 5; Mismatches 9; Indels 9; Gaps 0;

QY      1 ITPDQGGGQIGDNPILKAMKL 23
Db      147 VAPDYQGLGTIGAHPPYNAKSQL 169

RESULT 11
US-09-270-767-45060
; Sequence 45060, Application US/09270767
```

Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 45060
LENGTH: 502
TYPE: PRT
ORGANISM: *Drosophila melanogaster*
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45060

Query Match 35.9%; Score 46; DB 4; Length 502;
Best Local Similarity 56.2%; Pred. NO. 75;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 PDQGGGGQIGDNPPLK 18
Db 405 PDTGCGGAGGDRPSIK 420
||| ||| ||| :|

RESULT 12
US-08-568-459A-6
Sequence 6, Application US/08568459A
Patent No. 5843106
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 749 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum

US-08-568-459A-6
Query Match 35.9%; Score 46; DB 2; Length 749;
Best Local Similarity 47.1%; Pred. NO. 1.2e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGGGGQIGDNPPLK 18
Db 432 TPNSGGGNYGDRQISK 448
||| ||| ||| :|

RESULT 13
US-08-487-826B-6
Sequence 6, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 749 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-6

Query Match 35.9%; Score 46; DB 2; Length 749;
Best Local Similarity 47.1%; Pred. NO. 1.2e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGGGGQIGDNPPLK 18
Db 432 TPNSGGGNYGDRQISK 448
||| ||| ||| :|

RESULT 14
US-09-210-288-6
Sequence 6, Application US/09210288
Patent No. 6392026

```
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-09-210-288-6

Query Match 35.9%; Score 46; DB 3; Length 749;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 TPDGGGGGGGIGDNPILK 18
||:|||||:|
Db 432 TPNSSGGGNGYDGRQISK 448

RESULT 15
US-09-252-991A-20186
; Sequence 20186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20186
; LENGTH: 1098
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-20186

Query Match 35.9%; Score 46; DB 4; Length 1098;
Best Local Similarity 47.6%; Pred. No. 1.8e+02;
Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 1 ITPDGGGGGQ-----IGDNPIL 17
:|||||:|
Db 594 VQPDGGGAQRHRRAGEGPHL 614

RESULT 16
US-09-248-796A-22679
; Sequence 22679, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22679
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-22679

Query Match 35.2%; Score 45; DB 4; Length 82;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 5 GQGGGQIGDNPILKMLKLI 24
||:|||||:|
Db 27 GNGNGSGGNTILQILQLI 46

RESULT 17
US-09-543-681A-6484
; Sequence 6484, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6484
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-6484

Query Match 35.2%; Score 45; DB 4; Length 133;
Best Local Similarity 44.4%; Pred. No. 24;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 TPDGGGGGQIGDNPILKA 19
:|||||:|
Db 67 TTEVQGGVGVGEQPLIEA 84

RESULT 18
```

```

; OTHER INFORMATION: FUT1
US-09-151-592-2

Query Match      35.2%; Score 45; DB 3; Length 365;
Best Local Similarity 43.5%; Pred. NO. 75;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy      1 ITPDGGGGQIGDNPLLKAMLKL 23
      ||||:|:|:|:|:|:|
Db      83 IYPDGRFGNMGQYATLLALAQL 105

Search completed: January 25, 2005, 12:17:07
Job time : 31.3654 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:09:52 ; Search time 46.6346 Seconds
(without alignments)
193.681 Million cell updates/sec

Title: US-09-879-248-6_COPY_116_140

Perfect score: 128

Sequence: 1 ITPDGGGGGIGDNPLKXWLKLIJA 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pdb*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pdb*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pdb*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pdb*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pdb*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pdb*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pdb*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pdb*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pdb*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pdb*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pdb*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_NEW_PUB.pdb*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pdb*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	128	100.0	447	9	US-09-835-684-5
2	128	100.0	447	9	US-09-880-371-5
3	128	100.0	447	9	US-09-879-248-6
4	128	100.0	447	14	US-10-010-390-5
5	128	100.0	447	15	US-10-441-736-6
6	128	100.0	447	17	US-10-847-142-5
7	57	44.5	206	14	US-10-109-670-10
8	54	42.2	837	14	US-10-045-792-12
9	51	39.8	409	9	US-09-845-335-3
10	50	39.1	78	15	US-10-424-599-214915
11	50	39.1	177	16	US-10-437-963-149121
12	49	38.3	149	17	US-10-425-115-231795
13	49	38.3	483	14	US-10-369-493-3880
					Sequence 3880, Ap
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 10, Appl
					Sequence 3, Appli
					Sequence 214915,
					Sequence 149121,
					Sequence 231795,
					Sequence 3880, Ap

365	9	US-09-908-711-100	Sequence 100, App
43	14	US-10-269-806-163	Sequence 163, App
111	16	US-10-437-963-124136	Sequence 124136,
238	14	US-10-032-585-7297	Sequence 7297, Ap
207	17	US-10-425-115-301902	Sequence 301902,
81	17	US-10-425-115-347563	Sequence 347563,
102	15	US-10-424-598-216781	Sequence 216781,
134	17	US-10-425-115-351040	Sequence 351040,
213	17	US-10-425-115-210287	Sequence 210287,
435	14	US-10-369-493-6514	Sequence 6514, Ap
1343	9	US-09-815-242-11037	Sequence 11037, A
1343	15	US-10-282-122A-58196	Sequence 58196, A
353	15	US-10-374-780A-1704	Sequence 1704, Ap
354	16	US-10-437-963-194124	Sequence 194124,
354	16	US-10-437-963-195060	Sequence 195060,
408	16	US-10-781-014-462	Sequence 462, App
419	15	US-10-320-797-3176	Sequence 3176, Ap
441	9	US-09-738-626-5934	Sequence 5934, Ap
450	17	US-10-425-115-332466	Sequence 332466,
469	15	US-10-425-114-57241	Sequence 57241, A
749	13	US-10-153-273-6	Sequence 6, Appli
1086	9	US-09-924-154-15	Sequence 15, Appl
113	15	US-10-424-599-183658	Sequence 183658,
161	14	US-10-369-493-5035	Sequence 5035, Ap
880	16	US-10-437-963-119573	Sequence 119573,
23	14	US-10-269-806-176	Sequence 176, App
41	14	US-10-269-806-170	Sequence 182, App
46	14	US-10-269-806-182	Sequence 188, App
46	14	US-10-269-806-188	Sequence 128768,
69	17	US-10-437-963-128768	Sequence 128768,
79	17	US-10-425-115-333561	Sequence 333561,
104	17	US-10-425-115-239179	Sequence 239179,
108	17	US-10-425-115-220516	Sequence 220516,
113	15	US-10-264-049-2489	Sequence 2489, Ap
119	16	US-10-437-963-180990	Sequence 180990,
147	17	US-10-425-115-208583	Sequence 208583,
193	16	US-10-437-963-183514	Sequence 183514,
340	16	US-10-437-963-178172	Sequence 178172,
365	9	US-09-051-034A-4	Sequence 4, Appli
365	9	US-09-254-077A-10	Sequence 10, Appl
365	9	US-09-844-268-13	Sequence 13, Appl
365	9	US-09-844-705-13	Sequence 13, Appl
373	9	US-09-254-077A-12	Sequence 12, Appl
476	16	US-10-437-963-164622	Sequence 164622,
509	16	US-10-437-963-202306	Sequence 202306,
515	17	US-09-315-355-45	Sequence 46, Appl
518	17	US-10-425-115-239190	Sequence 239190,
546	15	US-10-425-114-60000	Sequence 60000, A
593	16	US-10-437-963-108680	Sequence 108680,
664	16	US-10-408-765A-2079	Sequence 2079, Ap
664	17	US-10-473-127-1658	Sequence 1658, Ap
692	16	US-10-437-963-161047	Sequence 161047,
789	15	US-10-282-122A-49250	Sequence 49250, A
874	15	US-10-282-122A-51045	Sequence 51045, A
159	17	US-10-425-115-305395	Sequence 305395,
453	15	US-10-425-114-49175	Sequence 49175, A
585	15	US-10-425-114-71739	Sequence 71739, A
655	17	US-10-425-115-305399	Sequence 305399,
701	14	US-10-156-761-12041	Sequence 12041, A
936	17	US-10-425-115-205479	Sequence 205479,
942	15	US-10-425-114-68187	Sequence 68187, A
45	17	US-10-425-115-34889	Sequence 324889,
51	17	US-10-425-115-235183	Sequence 235183,
95	15	US-10-425-114-52465	Sequence 52465, A
163	17	US-10-425-115-233374	Sequence 233374,
163	16	US-10-437-963-166199	Sequence 166199,
235	17	US-10-425-115-249433	Sequence 249433,
252	16	US-10-437-963-146435	Sequence 146435,
284	17	US-10-425-115-208441	Sequence 208441,
285	14	US-10-369-493-19127	Sequence 19127, A
285	16	US-10-437-963-185712	Sequence 185712,
389	14	US-10-369-493-4771	Sequence 4771, Ap
392	14	US-10-369-493-7531	Sequence 7531, Ap

87 44 34.4 394 14 US-10-369-493-1004 Sequence 1004, Ap
88 44 34.4 412 14 US-10-369-493-4993 Sequence 4993, Ap
89 44 34.4 439 16 US-10-437-963-166433 Sequence 166433, A
90 44 34.4 458 16 US-10-767-701-44941 Sequence 44941, A
91 44 34.4 686 14 US-10-032-585-7123 Sequence 7123, Ap
92 44 34.4 687 16 US-10-437-963-134108 Sequence 134108, A
93 44 34.4 774 15 US-10-282-122A-63166 Sequence 63166, A
94 44 34.4 875 15 US-10-282-122A-78522 Sequence 78522, A
95 44 34.4 890 14 US-10-156-761-14378 Sequence 14378, A
96 44 34.4 901 16 US-10-437-963-184559 Sequence 184559, A
97 43.5 34.0 78 17 US-10-425-115-339149 Sequence 339149, A
98 43.5 34.0 105 15 US-10-424-599-194061 Sequence 194061, A
99 43.5 34.0 120 17 US-10-425-115-209361 Sequence 209361, A
100 43.5 34.0 148 11 US-09-973-278-257 Sequence 257, App

ALIGNMENTS

RESULT 1
US-09-835-684-5
; Sequence 5, Application US/09835684
; Patent No. US20020019337A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/71
; CURRENT APPLICATION NUMBER: US/09/835,684
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-835-684-5

Query Match 100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLLKAMKLIA 25
Db 116 ITPDGGGGGQIGDNPLLKAMKLIA 140

RESULT 2
US-09-880-371-5
; Sequence 5, Application US/09880371
; Patent No. US20020059658A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: DeRoche, Jay
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 21829/91
; CURRENT APPLICATION NUMBER: US/09/880,371
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/211,585
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-880-371-5

Query Match 100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLLKAMKLIA 25
Db 116 ITPDGGGGGQIGDNPLLKAMKLIA 140

RESULT 3
US-09-879-248-6
; Sequence 6, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 21829/81
; CURRENT APPLICATION NUMBER: US/09/879,248
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/212,211
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-879-248-6

Query Match 100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLLKAMKLIA 25
Db 116 ITPDGGGGGQIGDNPLLKAMKLIA 140

RESULT 4
US-10-010-390-5
; Sequence 5, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/248,169
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-010-390-5

Query Match 100.0%; Score 128; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLLKAMKLIA 25
Db 116 ITPDGGGGGQIGDNPLLKAMKLIA 140

```

RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/203 (BEC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6

Query Match      100.0%; Score 128; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1      ITPDGGGGGQIGDNPPLKAMUKLIA 25
Db      116    ITPDGGGGGQIGDNPPLKAMUKLIA 140

RESULT 6
US-10-847-142-5
; Sequence 5, Application US/10847142
; Publication No. US20040265442A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Renick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSCE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; FILE REFERENCE: 21829/197
; CURRENT APPLICATION NUMBER: US/10/847,142
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/835,684
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-847-142-5

Query Match      100.0%; Score 128; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1      ITPDGGGGGQIGDNPPLKAMUKLIA 25
Db      116    ITPDGGGGGQIGDNPPLKAMUKLIA 140

RESULT 7
US-10-109-670-10
; Sequence 10, Application US/10109670
; Publication No. US20030105283A1
; GENERAL INFORMATION:

```



```

/ PRIOR APPLICATION NUMBER: US01/01320
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,853
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01349
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,902
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01239
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,870
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01348
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,882
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01347
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,896
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01307
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,864
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01341
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,856
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01336
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 09/764,868
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: US01/01312
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 60/179,065
/ PRIOR FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 60/180,628
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: 60/209,467
/ PRIOR FILING DATE: 2000-06-07
/ NUMBER OF SEQ ID NOS: 167
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 100
/ LENGTH: 365
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-908-711-100

Query Match 37.9%; Score 48.5
Best Local Similarity 54.5%; Pred. No.
Matches 12; Conservative 4; Mismatch

QY 2 TPQGQGGGTGDPNPLKAMKL 23
|||:|||||:
Db 284 TPDKGGQTQI-PNPLKNILRV 304

RESULT 15
US-10-269-806-163
/ Sequence 163, Application US/10269806
/ Publication No. US20030176352A1
/ GENERAL INFORMATION:
/ APPLICANT: Min, Hosung
/ APPLICANT: Stiney, Karen
/ APPLICANT: Hartley, Cynthia
/ TITLE OF INVENTION: Peptides and Related
/ FILE REFERENCE: A-750
/ CURRENT APPLICATION NUMBER: US/10/269,806
/ CURRENT FILING DATE: 2002-10-10
/ NUMBER OF SEQ ID NOS: 199
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 163
/ LENGTH: 43

```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized Peptide Sequence
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: At position 1, Fc at N-terminus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44)..(44)
; OTHER INFORMATION: At position 44, Fc at C-terminus
US-10-269-806-163
```

```
Query Match          37.5%; Score 48; DB 14; Length 43;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      5 GGGGGQIGDNPPLKAMLK 22
         |||||:|||||:
Db      21 GRGGGALRDGPTLKQWLE 38
```

```
RESULT 16
US-10-437-963-124136
; Sequence 124136, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437.963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124136
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(111)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_26904C.1.pep
US-10-437-963-124136
```

```
Query Match          37.5%; Score 48; DB 16; Length 111;
Best Local Similarity 55.6%; Pred. No. 36;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      5 GGGGGQIGDNPPLKAMLK 22
         |||||:|||||:
Db      4 GPGGGKIGKPLRVGLLK 21
```

```
RESULT 17
US-10-032-585-7297
; Sequence 7297, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
```

```
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7297
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7297
```

```
Query Match          37.5%; Score 48; DB 14; Length 238;
Best Local Similarity 43.5%; Pred. No. 82;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      2 TPDGGGGQIGDNPPLKAMLKLI 24
         |||||:|||||:
Db      50 TTDGNGGGGIGIVPTLQNIIVATV 72
```

```
RESULT 18
US-10-425-115-301902
; Sequence 301902, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 301902
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_3840C.1.pep
US-10-425-115-301902
```

```
Query Match          37.1%; Score 47.5; DB 17; Length 207;
Best Local Similarity 64.7%; Pred. No. 84;
Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
```

```
QY      5 GGGGGQIGDNPPLKAML 21
         |||||:|||||:
Db      77 GAGGGQDGDDE-LMRAML 92
```

```
RESULT 19
US-10-425-115-347563
; Sequence 347563, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 347563
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_8013C.1.pep
```

US-10-425-115-347563

Query Match 36.7%; Score 47; DB 17; Length 81;
 Best Local Similarity 40.0%; Pred. No. 36;
 Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPILKAM 20
 Db 17 VSPGALGKGQIGSRPVLNGL 36

RESULT 20

US-10-424-599-216781
 ; Sequence 216781, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 216781
 ; LENGTH: 102
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_37782C.1.pep
 US-10-424-599-216781

Query Match 36.7%; Score 47; DB 15; Length 102;
 Best Local Similarity 50.0%; Pred. No. 47;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 5 GGGGGQIGDNPILKAMLK 22
 Db 24 GGGGGQLGESFPWGEAFLK 41

Search completed: January 25, 2005, 12:18:58
 Job time : 51.6346 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:08:06 ; Search time 11.0577 Seconds
(without alignments)
217.534 Million cell updates/sec

Title: US-09-879-248-6_COPY_116_140
Perfect score: 128
Sequence: 1 ITPDGGGGQIGDNPLLKAWLKLI A 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR 79:*

1: piri: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	447	2	HrpW protein - Erw
2	54	42.2	837	2	DNA-Binding Protei
3	52	40.6	135	2	hemagglutinin homo
4	52	40.6	702	2	major surface prot
5	51	39.8	409	2	translation elonga
6	49.5	38.7	480	2	hypothetical prote
7	49	38.3	428	2	hypothetical prote
8	49	38.3	436	2	hypothetical prote
9	49	38.3	598	2	potassium channel
10	49	38.3	613	2	potassium channel
11	48	37.5	225	2	hypothetical prote
12	47	36.7	435	2	hypothetical prote
13	47	36.7	1343	1	DNA-directed RNA p
14	46.5	36.3	652	2	iron-regulated out
15	46.5	36.3	857	2	alanine-tRNA ligas
16	46	35.9	435	2	probable AAA-type
17	46	35.9	605	2	regulatory protein
18	46	35.9	1134	2	troponin C - Caeno
19	45.5	35.5	161	2	hypothetical prote
20	45.5	35.5	694	2	hypothetical prote
21	45	35.2	206	2	hypothetical 22K p
22	45	35.2	253	2	hypothetical prote
23	45	35.2	283	2	hypothetical prote
24	45	35.2	302	2	hypothetical prote
25	45	35.2	373	2	beta-galactoside a
26	45	35.2	405	2	hypothetical prote
27	45	35.2	405	2	molbdopterin bios
28	45	35.2	509	2	hypothetical prote
29	45	35.2	532	2	hypothetical prote

30	45	35.2	576	2	T17842	hypothetical prote
31	45	35.2	664	1	VERHOLA	lamin A - human
32	45	35.2	665	2	S27267	lamin A - rat
33	45	35.2	665	2	S28182	lamin A - mouse
34	45	35.2	1547	2	T13847	sno protein - frui
35	44	34.4	258	2	G72313	hypothetical prote
36	44	34.4	394	2	D69370	probable acyl-CoA
37	44	34.4	412	2	T32890	steroid/thyroid/re
38	44	34.4	501	2	A11031	probable membrane
39	44	34.4	535	2	S44827	F54P2.2 protein -
40	44	34.4	790	2	E40327	COI intron A prote
41	44	34.4	875	2	AE0401	alanine-tRNA ligas
42	44	34.4	890	2	P75289	alanyl-tRNA synthe
43	43.5	34.0	415	2	S60078	Runt domain contai
44	43.5	34.0	415	2	B55563	AML2a protein - hu
45	43.5	34.0	754	2	S41391	gelsolin - America
46	43.5	34.0	754	2	S53373	geloslin - America
47	43.5	34.0	962	2	AG2444	hypothetical prote
48	43	33.6	182	2	T18709	hypothetical prote
49	43	33.6	254	2	B72374	conserved hypotet
50	43	33.6	290	1	H70057	agmatinase homolog
51	43	33.6	328	2	E81452	probable fatty aci
52	43	33.6	341	2	AC3638	tetraacyldisacchar
53	43	33.6	410	2	G69274	probable acyl-CoA
54	43	33.6	414	1	JN0826	probable beta-keto
55	43	33.6	419	1	JC1211	beta-ketoacyl synt
56	43	33.6	437	2	F83854	hypothetical prote
57	43	33.6	454	2	AE1970	hypothetical prote
58	43	33.6	584	2	S62198	formylmethanofuran
59	43	33.6	597	2	S51212	BAK5 protein - bov
60	43	33.6	615	1	S08977	nif-specific regul
61	43	33.6	740	2	T03975	hypothetical prote
62	43	33.6	877	2	S76394	hypothetical prote
63	43	33.6	892	2	D69393	large helicase-rel
64	43	33.6	900	2	H90176	alanyl-tRNA synthe
65	42.5	33.2	239	2	S07377	ribosomal protein
66	42.5	33.2	253	1	R52PD4	ribosomal protein
67	42.5	33.2	265	2	H89834	ferriochrome transp
68	42.5	33.2	291	2	H71467	succinate-CoA liga
69	42.5	33.2	373	2	E72216	alanine-tRNA ligas
70	42.5	33.2	413	2	AC3614	tdp-4-dehydroham
71	42.5	33.2	652	2	S25265	outer membrane pro
72	42.5	33.2	717	2	S77190	hypothetical prote
73	42.5	33.2	860	2	A96717	unknown protein, 4
74	42	32.8	67	2	B35063	hypothetical prote
75	42	32.8	122	1	B69067	conserved hypotet
76	42	32.8	180	2	C86751	hypothetical prote
77	42	32.8	185	2	AD0834	hypothetical prote
78	42	32.8	233	1	DSBYN	superoxide dismuta
79	42	32.8	242	2	G81058	conserved hypotet
80	42	32.8	247	2	AG0250	conserved hypotet
81	42	32.8	260	1	A32515	acetolactate decar
82	42	32.8	266	2	S56520	hypothetical 29.1K
83	42	32.8	268	2	B64066	probable ABC trans
84	42	32.8	317	2	A64343	hypothetical prote
85	42	32.8	335	2	G82090	transcription regu
86	42	32.8	350	2	I55214	salivary proline-r
87	42	32.8	366	2	S77203	hypothetical prote
88	42	32.8	374	2	T36363	probable DNA-bindi
89	42	32.8	396	2	T01049	hypothetical prote
90	42	32.8	408	2	G87646	molbdopterin bios
91	42	32.8	409	2	H70303	ubiquinol-cytochro
92	42	32.8	459	2	C64679	phosphomannomutase
93	42	32.8	459	2	T41840	phosphohexosemutas
94	42	32.8	462	2	T42679	hypothetical prote
95	42	32.8	492	2	S28007	probable ATP-bindi
96	42	32.8	508	2	T09046	proline-rich prote
97	42	32.8	529	2	S12787	potassium channel
98	42	32.8	541	2	JX0112	signal recognition
99	42	32.8	546	2	T37665	probable t-complex
100	42	32.8	664	2	A56695	notch2 protein hom

ALIGNMENTS

RESULT 1
T18447
HrpW protein - Erwinia amylovora
C:Species: Erwinia amylovora
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18447
R:Gaudriault, S.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z18936
A:Accession: T18447
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-447 <GAU>
A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:e1293451; PIDN:CAA741
A:Experimental source: strain CFBP1430; specific host Pomoideae
C:Genetics:
A:Note: hrpW

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGQGGGQIGDNPPLKAMLKLI A 25
|||||
Db 116 ITPDGQGGGQIGDNPPLKAMLKLI A 140
|||||

RESULT 2
I57557
DNA-Binding Protein and transcription factor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: I57557
R:Quellie, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Cleve
Mol. Cell. Biol. 15, 3336-3343, 1995
A:Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosph
A:Reference number: I57557; MUID:95280934; PMID:7760829
A:Accession: I57557
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: mRNA
A:Residues: 1-837 <RES>
A:Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g10088
C:Genetics:
A:Gene: STAT6
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 42.2%; Score 54; DB 2; Length 837;
Best Local Similarity 64.3%; Pred. No. 5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGGQGGQIGDNPPLK 18
|||||
Db 802 GGGGSLGSGQLLK 815
|||||

RESULT 3
B49218
hemagglutinin homolog pmGA1.3 - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: B49218
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.
Infect. Immun. 61, 903-909, 1993
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut
A:Reference number: A49218; MUID:93162830; PMID:8432610
A:Accession: B49218
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-135 <MAR>
A:Cross-references: UNIPROT:Q53304; GB:S55216; NID:g265625; PIDN:AA825398.1; PID:g265625

F;174-176/Region: GTP-binding SAK/L motif

F;25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 39.8%; Score 51; DB 2; Length 409;
Best Local Similarity 57.1%; Pred. No. 6.5;
Matches 12; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 6 QGG--GGIGNPPLKAMKL 24
||| ||||| : |||:
DB 183 QGGASGGQGNPWVDKILKM 203

RESULT 6

hypothetical protein - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: E75433
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.O.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: E75433

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-480 <WHI>

A;Cross-references: UNIPROT:Q9RV87; GB:AE001963; GB:AE000513; NID:96458869; PIDN:AAF1071

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR142

A;Map position: 1

C;Superfamily: Deinococcus radiodurans hypothetical protein DR1142

Query Match 38.7%; Score 49.5; DB 2; Length 480;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 12; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 5 GGGGGGIGD-NPLLKAMKL 23
||||| : ||||| :
DB 157 GGGGGGIGDFAPIILAFRL 176

RESULT 7

hypothetical protein PH1701 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C;Accession: G71177
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekir
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: G71177

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-428 <KAW>

A;Cross-references: UNIPROT:O59360; GB:AP000007; NID:g3236134; PIDN:BAA30814.1; PID:g323

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH1701

Query Match 38.3%; Score 49; DB 2; Length 428;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 GGGIGNPPLKAMKL 23
||| ||||| : |||:
DB 286 GAQSGDNAMUKAMHRL 301

RESULT 8

T46107

hypothetical protein T25B15.140 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46107

R;Alcaraz, J.P.; Clabault, G.; Cottet, A.; Mache, R.; Mewes, H.W.; Lemcke, K.; Mayer, K.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23021

A;Accession: T46107

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-436 <ALC>

A;Cross-references: UNIPROT:Q9FT45; EMBL:AL132972

A;Experimental source: cultivar Columbia; BAC clone T25B15

C;Genetics:

A;Map position: 3

A;Introns: 418/1

A;Note: T25B15.140

Query Match 38.3%; Score 49; DB 2; Length 436;
Best Local Similarity 52.6%; Pred. No. 14;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPPLKAMKL 21
||| ||||| : |||:
DB 22 PDKPGSGQINSNSVLVAL 40

RESULT 9

S66669

potassium channel (Kv1.5) - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C;Accession: S66669

R;Sasaki, Y.; Ishii, K.; Nunoki, K.; Yamagishi, T.; Taira, N.

FEBS Lett. 372, 20-24, 1995

A;Title: The voltage-dependent K(+) channel (Kv1.5) cloned from rabbit heart and facilitat

A;Reference number: S66669; MUID:96032538; PMID:7556635

A;Accession: S66669

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-598 <SAS>

A;Cross-references: UNIPROT:P50638; EMBL:D45025; NID:g1060972; PIDN:BAA08082.1; PID:d1006

C;Superfamily: potassium channel protein drkl

Query Match 38.3%; Score 49; DB 2; Length 598;
Best Local Similarity 45.5%; Pred. No. 20;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 PDGGGGGQIGDNPPLKAMKL 24
||| ||||| : |||:
DB 366 PGGGGGGGQGNQMSLILRLV 387

RESULT 10

A56031

potassium channel KCNA5 - human

N;Alternate names: potassium channel HK2; potassium channel PCN1; shaker-related potassi

C;Species: Homo sapiens (man)

C;Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004

C;Accession: A56031; A38556; B3922; A38074

R;Philipson, L.H.; LaMendola, J.; Bell, G.I.; Steiner, D.F.

submitted to GenBank, September 1990

A;Reference number: A56031

A;Accession: A56031

A;Molecule type: mRNA

A;Residues: 1-613 <PHI>

A;Cross-references: UNIPROT:P22460; GB:M55513; NID:g189653; PIDN:AAA36422.1; PID:g189654

R;Philipson, L.H.; Hice, R.E.; Schaefer, K.; LaMendola, J.; Bell, G.I.; Neelson, D.J.; St

Proc. Natl. Acad. Sci. U.S.A. 88, 53-57, 1991

A;Title: Sequence and functional expression in Xenopus oocytes of a human insulinoma and

A:Reference number: A38556; MUID:91095456; PMID:1986382
A:Accession: A38556
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-56, 'G', 58-137, 'L', 139-213, 'R', 215-227, 'P', 229-545, 'QG', 546-613 <PH2>
A:Cross-references: GB:M55513
R: Tamkun, M.M.; Knoch, K.M.; Walbridge, J.A.; Kroemer, H.; Roden, D.M.; Glover, D.M.
FASEB J. 5, 331-337, 1991
A:Title: Molecular cloning and characterization of two voltage-gated K(+) channel cDNAs
A:Reference number: A39922; MUID:91160866; PMID:2001794
A:Accession: B39922
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-54, 56-137, 'L', 139-186, 'G', 189-213, 'R', 215-227, 'P', 229-297, 'PQQRGH', 309-558
A:Cross-references: GB:M60451
R: Curran, M.E.; Landes, G.M.; Keating, M.T.
Genomics 12, 729-737, 1992
A:Title: Molecular cloning, characterization, and genomic localization of a human potassium channel cDNA
A:Reference number: A38074; MUID:92241872; PMID:1349297
A:Accession: A38074
A:Molecule type: DNA
A:Residues: 1-137, 'L', 139-153, 'R', 155-213, 'R', 215-227, 'P', 229-281, 'V', 283-578, 'QLPPREV'
A:Cross-references: GB:M83254; NID:9190202; PIDN:AAA60146.1; PID:9190203
A:Experimental source: heart
A:Note: sequence extracted from NCBI backbone (NCBI:98573, NCBIP:98577)
C:Genetics:
A:Gene: GDB:KCNAB5
A:Cross-references: GDB:127904; OMIM:176267
A:Map position: 12p13.33-12p13.31
C:Superfamily: potassium channel protein drkl
C:Keywords: glycoprotein; phosphoprotein; potassium channel; transmembrane protein; voltage-gated; ion channel; binding site: phosphate (Ser) (covalent) #status predicted
F:557/Binding site: phosphate (Ser) (covalent) #status predicted
Query Match 38.3%; Score 49; DB 2; Length 613;
Best Local Similarity 45.5%; Pred. No. 21;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 3 PDQGGGGGQIGDNPPLKAMKLI 24
DB 381 PGGGGGGGQGGQMSLALIRVI 402
RESULT 11
AF2140
hypothetical protein all2677 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2140
R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2140
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-225 <KUR>
A:Cross-references: UNIPROT:Q8VTN7; GB:BA000019; PIDN:BA074376.1; PID:917131770; GSPDB:G000000001
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all2677

Query Match 37.5%; Score 48; DB 2; Length 225;
Best Local Similarity 43.5%; Pred. No. 9.7;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 3 PDQGGGGGQIGDNPPLKAMKLI 25
DB 137 PEATGGLIGPTLTPHPQLVA 159

RESULT 12
T20152
hypothetical protein C52E4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20152
R: Holt, R.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19230
A:Accession: T20152
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-435 <WIL>
A:Cross-references: UNIPROT:Q18787; EMBL:Z78012; PIDN:CAB01414.1; GSPDB:GN00023; CESP:C52E4
A:Experimental source: clone C52E4
C:Genetics:
A:Gene: CESP:C52E4.4
A:Map position: 5
A:Introns: 26/1; 60/3; 137/3; 318/3
C:Superfamily: ATP-dependent 26S proteinase; PtsH/SEC18/CDC48-type ATP-binding domain homolog
Query Match 36.7%; Score 47; DB 2; Length 435;
Best Local Similarity 52.4%; Pred. No. 29;
Matches 11; Conservative 3; Mismatches 3; Indels 4; Gaps 1;
QY 4 DQGGGGGQIGDNPPLKAMKLI 24
DB 289 DQGGGGGQIGDNPPLKAMKLI 305
RESULT 13
H64073
DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Haemophilus influenzae (strain Rd 1)
C:Species: Haemophilus influenzae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: H64073
R: Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: H64073
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1343 <TIGR>
A:Cross-references: UNIPROT:P43738; GB:U32733; GB:L42023; NID:91573493; PIDN:AAC22173.1;
C:Superfamily: DNA-directed RNA polymerase beta chain
C:Keywords: nucleotidyltransferase; transcription
Query Match 36.7%; Score 47; DB 1; Length 1343;
Best Local Similarity 36.4%; Pred. No. 18+02;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 3 PDQGGGGGQIGDNPPLKAMKLI 24
DB 412 PEGEGAGILSNEDIIAVWRKLI 433
RESULT 14
D82317
iron-regulated outer membrane virulence protein, TonB receptor family VC0475 [imported]
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82317
R: Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Church, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82317

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-652 <HEI>
A;Cross-references: UNIPROT:P27772; GB:AE004134; GB:AE003852; NID:G9654900; PIDN:AAF9364
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0475
A;Map position: 1
C;Superfamily: ferrienterochelin receptor; tonB-dependent receptor amino-terminal homolo

Query Match 36.3%; Score 46.5; DB 2; Length 652;
Best Local Similarity 61.9%; Pred. No. 54;
Matches 13; Conservative 1; Mismatches 4; Indels 3; Gaps 2;

Qy 1 ITPD-GQ--GGGQIGDNPPLK 18
:|||||:|||||:|||||
Db 428 VTPDWGQVSGGNIYGNPDLK 448

RESULT 15
H70411
A;Title: alanine-trna ligase (EC 6.1.1.7) - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: H70411
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
C;Accession: H70411
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-867 <AQF>
A;Cross-references: UNIPROT:O67323; GB:AE000733; NID:G2983720; PIDN:AAC07289.1; PID:G298
A;Experimental source: strain VF5
C;Genetics:
A;Gene: alas
C;Superfamily: alanyl-trna ligase
C;Keywords: aminoacyl-trna synthetase; ligase; protein biosynthesis

Query Match 36.3%; Score 46.5; DB 2; Length 867;
Best Local Similarity 47.6%; Pred. No. 73;
Matches 10; Conservative 5; Mismatches 3; Indels 3; Gaps 1;

Qy 6 QGGGQIGDNPPL---KAMKL 23
:|||||: :|||:
Db 495 EGGGQIGDAGIIESDKALFKV 515

RESULT 16
F84674
A;Title: probable AAA-type ATPase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84674
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
C;Accession: F84674
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-435 <STO>
A;Cross-references: UNIPROT:Q9ZNT0; GB:AE002093; NID:G3860272; PIDN:AAC73040.1; GSPDB:GN
C;Genetics:
A;Gene: At2g27600
A;Map position: 2

Query Match 35.9%; Score 46; DB 2; Length 435;
Best Local Similarity 58.3%; Pred. No. 41;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DGGGQIGDNP 15
|||:|:|:|:
Db 98 DGGGQIGDNP 109

RESULT 17
S01066
A;Title: regulatory protein nifa - Bradyrhizobium japonicum
C;Species: Bradyrhizobium japonicum
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 28-Apr-2003
C;Accession: S01066
R;Hoehn, B.; Fischer, H.M.; Anthamatten, D.; Bruderer, T.; Hennecke, H.
Nucleic Acids Res. 15, 8479-8499, 1987
A;Title: The symbiotic nitrogen fixation regulatory operon (fixRnifa) of Bradyrhizobium
A;Reference number: S01065; MUID:88040468; PMID:3313281
C;Accession: S01066
A;Molecule type: DNA
A;Residues: 1-605 <THO>
A;Cross-references: EMBL:X06167; NID:G39526; PIDN:CAA29531.1; PID:G39528
C;Genetics:
A;Gene: nifa
C;Superfamily: Response regulator (sigma54-dependent transcriptional activator), Fh1A
F;253-474/Domain: RNA polymerase sigma factor interaction domain homology <SFI>

Query Match 35.9%; Score 46; DB 2; Length 605;
Best Local Similarity 52.6%; Pred. No. 59;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 6 QGGGQIGDNPPLKAMKL 24
|:|:|:|:|:|:|:
Db 249 QAHGIIGDSPALSALLEKI 267

RESULT 18
D75014
A;Title: hypothetical protein PAB1257 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: D75014
R;anonymous, Genoscope
Submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc
A;Reference number: A75001
A;Accession: D75014
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1134 <KAW>
A;Cross-references: UNIPROT:Q9UY58; GB:AJ248288; GB:AL096836; NID:G5458960; PIDN:CAB5055/
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1257

Query Match 35.9%; Score 46; DB 2; Length 1134;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 10 QIGDNPPLKAMKL 22
|:|:|:|:|:|:|:
Db 646 QVGDNPFLFSEILK 658

RESULT 19
T42006
A;Title: troponin C - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T42006; T41897
R;Terami, H.; Williams, B.D.; Kitamura, S.; Sakube, Y.; Matsumoto, S.; Doi, S.; Obinata,
J. Cell Biol. 146, 193-202, 1999
A;Title: Genomic organization, expression, and analysis of the troponin C gene pat-10 of
A;Reference number: Z22021; MUID:99332308; PMID:10402470

Search completed: January 25, 2005, 12:15:59
Job time : 15.0577 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:07:26 ; Search time 94.2308 Seconds
(without alignments)
152.650 Million cell updates/sec

Title: US-09-879-248-6_COPY_116_140
Perfect score: 128
Sequence: 1 ITPDGGGGGIGDNPLKAWLKLIA 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database : Uniprot 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	128	100.0	447	2	O54508 erwinia amy
2	128	100.0	447	2	O9LAW2
3	128	100.0	450	2	O6XDB8
4	128	100.0	450	2	AAQ17046
5	128	100.0	450	2	AA545453
6	120	93.8	138	2	O79AW7
7	71	55.5	479	2	O6RK52
8	71	55.5	479	2	O6D5C8
9	71	55.5	479	2	AA520352
10	60	46.9	237	2	O873M4
11	57	44.5	174	2	O9P923
12	56	43.8	165	2	O9P920
13	54	42.2	172	2	O9P921
14	54	42.2	220	2	O74200
15	54	42.2	837	1	STA6 MOUSE
16	53	41.4	173	2	O9P919
17	52	40.6	133	2	O49469
18	52	40.6	135	2	O53304
19	52	40.6	210	1	SODM_ASPFU
20	52	40.6	229	2	O6CFAL
21	52	40.6	584	2	O9KH13
22	52	40.6	692	2	O7NAN9
23	52	40.6	702	2	O49498
24	52	40.6	708	2	O7NAP2
25	52	40.6	710	2	O7NBR1
26	51	39.8	409	1	EFTU SYNTP6
27	51	39.8	409	1	EFTU SYNTP7
28	51	39.8	578	2	O7S3B9
29	51	39.8	667	2	O7NBR3
30	51	39.8	697	2	O7NBS5
31	50.5	39.5	504	2	O6BXD4

32	50.5	39.5	891	2	O7VQG3
33	50	39.1	61	2	P70383
34	50	39.1	177	2	O9FTW3
35	50	39.1	889	1	SYA_SYNXP
36	50	39.1	892	2	O7VEG5
37	50	39.1	1759	2	O73XY5
38	50	39.1	1759	2	AA504490
39	49.5	38.7	97	2	O8LOW7
40	49.5	38.7	480	2	O9RV87
41	49.5	38.7	962	2	O8DIY4
42	49	38.3	428	2	O59360
43	49	38.3	436	2	O9FT45
44	49	38.3	452	2	O7RV88
45	49	38.3	593	2	O28248
46	49	38.3	598	1	CIK5_RABIT
47	49	38.3	600	2	O866G0
48	49	38.3	601	1	CIK5_MUSPF
49	49	38.3	613	1	CIK5_HUMAN
50	48.5	37.9	381	2	O7Z6M4
51	48.5	37.9	381	2	AAH53533
52	48.5	37.9	581	2	O9FLE2
53	48.5	37.9	657	2	O8IR65
54	48.5	37.9	671	2	O6RX11
55	48.5	37.9	671	2	AA548941
56	48	37.5	172	2	O8P4W4
57	48	37.5	225	2	O8YTN7
58	48	37.5	238	1	TBP CANAL
59	48	37.5	273	2	O75135
60	48	37.5	273	2	AA887235
61	48	37.5	321	2	O6RYW6
62	48	37.5	321	2	AA883904
63	48	37.5	412	2	O6VNR9
64	48	37.5	412	2	AAQ63913
65	48	37.5	505	2	O9V9Y0
66	48	37.5	548	2	O8U328
67	48	37.5	904	1	SYA_SULTO
68	48	37.5	1831	2	O9W2E1
69	48	37.5	1872	2	O7YU83
70	47.5	37.1	262	2	O89PD9
71	47	36.7	316	2	O9XFK6
72	47	36.7	355	2	O6EV72
73	47	36.7	388	2	O8IIM8
74	47	36.7	425	2	O748W0
75	47	36.7	425	2	AAK36283
76	47	36.7	435	1	PR57_CAEEL
77	47	36.7	443	2	P79770
78	47	36.7	449	2	O8XW11
79	47	36.7	737	2	O8RUT8
80	47	36.7	872	2	O9VAMS
81	47	36.7	874	2	O8T9B8
82	47	36.7	901	2	O89RG3
83	47	36.7	1343	1	RPOB_HAEIN
84	46.5	36.3	353	2	O84R60
85	46.5	36.3	391	2	O7S2S2
86	46.5	36.3	391	2	BAD09637
87	46.5	36.3	504	2	O9V9L8
88	46.5	36.3	652	1	IRGA_VIBCH
89	46.5	36.3	867	1	SYA_AQUAE
90	46	35.9	55	2	O8U5V1
91	46	35.9	173	2	O79YP3
92	46	35.9	173	2	O8R8U6
93	46	35.9	182	2	O855V7
94	46	35.9	288	2	O866E3
95	46	35.9	290	2	O866E2
96	46	35.9	366	2	O866B9
97	46	35.9	366	2	O866E1
98	46	35.9	411	2	O8KUX9
99	46	35.9	434	2	O8LAK9
100	46	35.9	435	2	O9ZNT0

ALIGNMENTS

```

RESULT 1
OS4508
ID O54508 PRELIMINARY; PRT; 447 AA.
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW protein.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ea246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSSP; Q9RHW0; IE66.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 116 ITPDGGGGGQIGDNPLKMLKLLIA 140

RESULT 2
OS4508
ID O54508 PRELIMINARY; PRT; 447 AA.
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW protein.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFBP1430;
RA Gaudriault S., Brisset M.N., Barny M.A.;
RL "DAPA, an essential pathogenicity factor of Erwinia amylovora showing homology with AvrB of Pseudomonas syringae, is secreted via the Hrp secretion pathway in a DspB-dependent way.";
RL Mol. Microbiol. 26:1057-1069(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CFBP1430;
RA Gaudriault S., Brisset M.N., Barny M.A.;
RL "HrpW of Erwinia amylovora, a new Hrp-secreted protein.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Ea321;
RA Kim J.F., Zmoff C.H., Beer S.V.;
RL "HrpW, a new harpin of Erwinia amylovora, is a member of a family of pectate lyases.";
RL Phycopathology 87:0-0(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Ea321;
RA Kim J.F., Beer S.V.;
RL "HrpW of Erwinia amylovora, a new harpin that contains a domain homologous to pectate lyases of a distinct class.";
RL J. Bacteriol. 180:5203-5210(1998).
DR EMBL; Y13831; CAA74158.1; -.
DR EMBL; U94513; AAC62314.1; -.
DR PIR; T18447; T18447.
DR HSSP; Q9RHW0; IE66.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 116 ITPDGGGGGQIGDNPLKMLKLLIA 140

RESULT 2
OS4508
ID O54508 PRELIMINARY; PRT; 447 AA.
DT 01-JUN-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Harpin HrpW.

```

```

GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ea246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSSP; Q9RHW0; IE66.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45340 MW; 0BBAA3871EDC2F6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 5.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 116 ITPDGGGGGQIGDNPLKMLKLLIA 140

RESULT 3
OS4508
ID O6XDB8 PRELIMINARY; PRT; 450 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K., Hur J.H., Lim C.K.;
RA Hur J.H., Lim C.K.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I., Hur J.H., Lim C.K.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
DR EMBL; AY330755; AAS45453.1; -.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGQIGDNPLKMLKLLIA 25
Db 114 ITPDGGGGGQIGDNPLKMLKLLIA 138

RESULT 4
AAQ17046 PRELIMINARY; PRT; 450 AA.
ID AAQ17046
AC AAQ17046;

```

```
DT 03-MAR-2004 (TrEMBLrel. 27, Created)
DT 03-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 03-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shretha R.; Baek S.J.; Bae H.N.; Cho J.M.; Park D.H.; Hwang I.K.;
RA Hur J.H.; Lim C.K.;
RT "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RT and expression of hrpG encoding elicitor of the hypersensitive
RT response.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY237642; AAQ17046.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGIGDNPLLKAMKLIA 25
Db 114 ITPDGGGGGIGDNPLLKAMKLIA 138

RESULT 5
AAS45453
ID AAS45453 PRELIMINARY; PRT; 450 AA.
AC AAS45453
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT3;
RA Shretha R.; Baek S.J.; Park D.H.; Bae H.N.; Cho J.M.; Hwang I.;
RA Hur J.H.; Lim C.K.;
RT "Identification of hrp genes cluster and characterization of HR
RT elicitor hrpNep gene in Erwinia pyrifoliae.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY530755; AAS45453.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGIGDNPLLKAMKLIA 25
Db 114 ITPDGGGGGIGDNPLLKAMKLIA 138

RESULT 6
Q79AW7
ID Q79AW7 PRELIMINARY; PRT; 138 AA.
AC Q79AW7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW (Fragment).
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX NCBI_TaxID=218491;
```

```
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Es321;
RX MEDLINE=98115919; PubMed=9448330;
RA Bogdanove A.J.; Kim J.F.; Wei Z.; Kolchinsky P.; Charkowski A.O.;
RA Conlin A.K.; Collmer A.; Beer S.V.;
RT "Homology and functional similarity of an hrp-linked pathogenicity
RT locus, dsrEF, of Erwinia amylovora and the avirulence locus avrE of
RT Pseudomonas syringae pathovar tomato.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330(1998).
DR EMBL; U97504; AAC04849.1; -.
FT NON_TER 138
SQ SEQUENCE 138 AA; 13788 MW; 4FE177177C74B3C6 CRC64;

Query Match 93.8%; Score 120; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ITPDGGGGGIGDNPLLKAMKL 23
Db 116 ITPDGGGGGIGDNPLLKAMKL 138

RESULT 7
Q6RK52
ID Q6RK52 PRELIMINARY; PRT; 479 AA.
AC Q6RK52
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
OS Pectobacterium atrosepticum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1039;
RA Holveva M.C.; Bell K.S.; Hyman L.J.; Avrova A.O.; Whisson S.C.;
RA Birch P.R.J.; Toth I.K.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY496066; AAS20352.1; -.
DR InterPro; IPR011050; Pectin lyase-like.
DR InterPro; IPR004898; Pect. lyase.
DR Pfam; PF03211; Pectate lyase, 1.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 ITPDGGGGGIGDNPLLKAMKL 24
Db 166 INPTADGGGQLSGNDLLKALLEI 189

RESULT 8
Q6D5C8
ID Q6D5C8 PRELIMINARY; PRT; 479 AA.
AC Q6D5C8
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Type III effector protein.
GN Name=hrpW; ORFNames=ECA2112;
OS Erwinia carotovora subsp. atroseptica SCRI1043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
```

```

RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1043;
RA Bell K.S., Sebailhia M., Pritchard L., Holden M., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmund G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.,
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX950851; CAG75014.1; -.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGIGDNPPLKMLKLI 24
Db 166 INPTADGGGQLSGNDLLKALLELI 189

RESULT 9
AAS20352
ID AAS20352 PRELIMINARY; PRT; 479 AA.
AC AAS20352;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE HrpW.
GN HrpW.
OS Pectobacterium atrosepticum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RA Holeva M.C., Bell K.S., Hyman L.J., Avrova A.O., Whisson S.C.,
RA Birch P.R.J., Toth I.K.;
RT "Rapid demonstration of a role early in disease development for the
RT type III secretion system of Erwinia carotovora subsp. atroseptica
RT SCRI1039 using a pooled transposon mutation grid.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY496066; AAS20352.1; -.
SQ SEQUENCE 479 AA; 46706 MW; 3463E226CDF0406A CRC64;

Query Match 55.5%; Score 71; DB 2; Length 479;
Best Local Similarity 58.3%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGIGDNPPLKMLKLI 24
Db 166 INPTADGGGQLSGNDLLKALLELI 189

RESULT 10
Q873M4
ID Q873M4 PRELIMINARY; PRT; 237 AA.
AC Q873M4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Manganese superoxide dismutase (fragment).
GN Name=MnSOD;
OS Malassezia sympodialis.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Exobasidiomycetidae; Malasseziales; Malassezia.
OX NCBI_TaxID=76777;
RN [1]
RP SEQUENCE FROM N.A.
RA Andersson A., Rasool O., Schmidt M., Kodzius R., Crameri R.,
RA Scheynius A.;

```

```

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AJ548421; CAD68071.1; -.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase; 1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 237 AA; 26716 MW; 6C2096462E8402E9 CRC64;

Query Match 46.9%; Score 60; DB 2; Length 237;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ITPDGGGGGIGDNPPLKMLK 22
Db 118 MAPQSGGGQLNDGFLKQAIDK 139

RESULT 11
Q9P923
ID Q9P923 PRELIMINARY; PRT; 174 AA.
AC Q9P923;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Manganese-dependent superoxide dismutase (EC 1.15.1.1)
DE (fragment).
GN Name=soda;
OS Pneumocystis carinii f. sp. muris.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=42066;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20427437; PubMed=10975697;
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferragut C., Viscogliosi E.,
RA Dei-Cas E., Wakefield A.E.;
RT "Genetic divergence at the SODA locus of six different formae
RT speciales of Pneumocystis carinii.";
RL Med. Mycol. 38:289-300(2000).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF146751; AAF25723.1; -.
DR HSSP; Q92450; IKKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 174 AA; 20198 MW; 2730D8F435576124 CRC64;

```

OC	Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC	Pneumocystis.
OX	NCBI_TaxID=42068;
[1]	SEQUENCE FROM N.A.
RP	MEDLINE=20427437; PubMed=10975697;
RX	Denis C.M., Mazare E., Guyot K., Odberg-Ferragut C., Viscogliosi E.,
RA	Del-Cas E., Wazefield A.E.;
RA	"Genetic divergence at the SODA locus of six different formae
RT	speciales of Pneumocystis carinii.";
RT	Med. Mycol. 38:289-300(2000).
RL	CC -i- FUNCTION: Destroys radicals which are normally produced within the
CC	cells and which are toxic to biological systems (By similarity).
CC	-i- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC	-i- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC	family.
CC	ENBL; AF146753; AAF25725.1; -.
DR	HSP; Q92450; 1KKC.
DR	GO; GO:0046872; P:metal ion binding; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR	GO; GO:0006801; P:superoxide metabolism; IEA.
DR	InterPro; IPR001189; SODismutase.
DR	Pfam; PF02777; Sod_Fe_C; 1.
DR	Pfam; PF00081; Sod_Fe_N; 1.
DR	PRINTS; PR01703; MNSODISMUTASE.
DR	ProDom; PD000475; SODismutase; 1.
DR	PROSITE; PS00088; SOD_MW; 1.
KW	Oxidoreductase.
FT	NON_TER 1 1
FT	NON_TER 172 172
SQ	SEQUENCE 172 AA; 19884 MW; CDEC754E7F78F2A CRC64;
	Query Match 42.2%; Score 54; DB 2; Length 172;
	Best Local Similarity 45.5%; Pred. No. 12;
	Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps
Qy	1 ITPDGGGGGIGDNPFLKAMLK 22
	: : : : : : : : : :
Dd	69 LLPPKGGGGQVIDGLPLVDAIKK 90
RESULT 14	
O74200	PRELIMINARY; PRT; 220 AA.
ID	O74200
AC	O74200;
DT	01-NOV-1998 (TrEMBLrel. 08, Created)
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Manganese superoxide dismutase precursor.
GN	Name=mnsood;
OS	Pneumocystis carinii.
OC	Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC	Pneumocystis.
OX	NCBI_TaxID=4754;
[1]	SEQUENCE FROM N.A.
RP	Denis C.M., Guyot K., Del-Cas E., Camus D., Odberg-Ferragut C.;
RA	Submitted (DEC-1997) to the ENBL/GenBank/DBJ databases
RL	CC -i- FUNCTION: Destroys radicals which are normally produced within the
CC	cells and which are toxic to biological systems (By similarity).
CC	-i- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC	-i- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC	family.
CC	ENBL; AF036321; AAC24764.1; -.
DR	HSP; Q92450; 1KKC.
DR	GO; GO:0046872; P:metal ion binding; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR	GO; GO:0006801; P:superoxide metabolism; IEA.
DR	InterPro; IPR000834; Peptidase_M14.
DR	InterPro; IPR001189; SODismutase.
DR	Pfam; PF02777; Sod_Fe_C; 1.


```
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase; Signal.
FT SIGNAL 1 27 Potential.
FT CHAIN 48 169 manganese superoxide dismutase.
FT CHAIN 170 183 manganese superoxide dismutase.
FT CHAIN 184 207 manganese superoxide dismutase.
FT CHAIN 208 220 manganese superoxide dismutase.
FT CHAIN 28 35 manganese superoxide dismutase.
FT CHAIN 36 47 manganese superoxide dismutase.
SQ SEQUENCE 220 AA; 25869 MM; 73B1F11C98929E18 CRC64;

Query Match 42.2%; Score 54; DB 2; Length 220;
Best Local Similarity 45.5%; Pred. No. 16;
Matches 10; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 ITPDGGGGGQIGDNPFLKMLK 22
: | : ||||| : ||| : |
Db 106 LLPNRQGGGEIINGELVEAIKK 127

RESULT 15
STA6 MOUSE STANDARD; PRT; 837 AA.
ID STA6 MOUSE STANDARD; PRT; 837 AA.
AC P52833;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and transcription activator 6.
GN Name=Stat6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95280934; PubMed=7760829;
RA Quelle F.W., Shmida K., Thierfelder W., Fischer C.L., Kim A.,
RA Ruben S.M., Cleveland J.L., Pierce J.H., Keegan A.D., Nelms K.,
RA Paul W.E., Ihle J.N.;
RT "Cloning of murine Stat6 and human Stat6, Stat proteins that are
RT tyrosine phosphorylated in responses to IL-4 and IL-3 but are not
RT required for mitogenesis";
RL Mol. Cell. Biol. 15:3336-3343(1995).
CC -!- FUNCTION: Carries out a dual function: signal transduction and
CC activation of transcription. Involved in interleukin-4 signalling.
CC -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC member (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC in response to phosphorylation.
CC -!- PTM: Tyrosine phosphorylated following stimulation by IL-4 and IL-
CC 3.
CC -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L47650; AAT79006.1; -.
DR F1R; I57557; I57557.
DR HSSP; P42227; IBGI.
DR MGD; MGI:103034; Stat6.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0042127; P:regulation of cell proliferation; IDA.
```

```
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
KW Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW Transcription regulation.
FT DOMAIN 517 632 SH2.
FT MOD RES 641 641 Phosphotyrosine (by JAK) (By similarity).
SQ SEQUENCE 837 AA; 93725 MM; FD7D7C448743EFC7 CRC64;

Query Match 42.2%; Score 54; DB 1; Length 837;
Best Local Similarity 64.3%; Pred. No. 64;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGGGQIGDNPFLK 18
| : ||| : |||
Db 802 GGGGSLGSQLK 815

RESULT 16
Q9P919 PRELIMINARY; PRT; 173 AA.
ID Q9P919 PRELIMINARY; PRT; 173 AA.
AC Q9P919;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Manganese-dependent superoxide dismutase (EC 1.15.1.1)
DE (Fragment).
DE Name=sodA;
OS Pneumocystis carinii f. sp. suis.
OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC Pneumocystis.
OX NCBI_TaxID=112251;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20427437; PubMed=10975697;
RA Denis C.M., Mazars E., Guyot K., Odberg-Ferragut C., Viscogliosi E.,
RA Del-Cas E., Wakefield A.E.;
RT "Genetic divergence at the SODA locus of six different formae
RT specialises of Pneumocystis carinii.";
RL Med. Mycol. 38:289-300(2000).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR EMBL; AF146755; AAF25727.1; -.
DR HSSP; Q92450; 1KKC.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004784; F:superoxide dismutase activity; IEA.
DR GO; GO:0006801; P:superoxide metabolism; IEA.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF02777; Sod_Fe_C; 1.
DR Pfam; PF00081; Sod_Fe_N; 1.
DR PRINTS; PR01703; MNSODISMUTASE.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON TER 1
FT NON TER 173
SQ SEQUENCE 173 AA; 20099 MM; 69110EFD0005665B CRC64;

Query Match 41.4%; Score 53; DB 2; Length 173;
Best Local Similarity 40.9%; Pred. No. 17;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 ITPDGGGGQIGDNPFLKMLK 22
```


Qy

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:06:11 ; Search time 60.75 Seconds
(without alignments)
159.435 Million cell updates/sec

Title: us-09-879-248-6_copy_31_57
Perfect score: 128
Sequence: 1 NSALGQPIDRQTIEQMAQLLAKLS 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	2	AAW96260 Hypersens
2	128	100.0	447	3	AAW71094 Erwinia a
3	128	100.0	447	3	AAW84855 A hyperse
4	128	100.0	447	5	AAO22548 Hyperse
5	128	100.0	447	5	AAE18296 Erwinia a
6	128	100.0	447	5	AAE16448 E. amylov
7	128	100.0	447	5	ABO92226 Erwinia a
8	52	40.6	313	7	ABO73754 Pseudomon
9	52	40.6	457	7	ABO73936 Pseudomon
10	52	40.6	710	8	ADN73161 Thale cre
11	51	39.8	660	7	ADF29258 Bacterial
12	51	39.8	660	7	ADG32467 Pasteurel
13	51	39.8	1312	2	AAW22775 Human RAD
14	51	39.8	1312	2	AAW71295 Human hom
15	51	39.8	1318	5	ABW77985 Amino aci
16	51	39.8	1318	7	ADJ68860 Human hea
17	51	39.8	1318	8	ADJ65509 RAD50 hom
18	50.5	39.5	286	5	ABG70071 Human pre
19	50.5	39.5	352	2	AAW22307 Lipase mo
20	50.5	39.5	596	5	ABG61825 Prostate
21	50.5	39.5	596	7	ABW75398 Prostate
22	50	39.1	430	5	ABW47515 Listeria
23	49	38.3	264	7	ABO71596 Pseudomon
24	49	38.3	697	6	ABR40712 Oryza sat
25	49	38.3	3249	6	ABU39648 Protein e

99 44.5 34.8 859 7 ADC00051 Enterohae
100 44.5 34.8 1153 6 ABU33426 Protein e

ALIGNMENTS

RESULT 1
AAW96260
ID AAW96260 standard; protein; 447 AA.
AC AAW96260;
XX
DT 14-JUN-1999 (first entry)
XX
DE Hypersensitive response eliciting protein (HrpW).
XX
KW Hypersensitive response elicitor protein; hypersensitive response; hrpW;
KW pathogen; infection; crop protection; disease resistance;
KW pest resistance; transgenic plant; colouration; maturation.
XX
OS Erwinia amylovora.
XX
FN WO9907208-A1.
XX
PD 18-FEB-1999.
XX
PF 27-JUL-1998; 98WO-US015547.
XX
PR 06-AUG-1997; 97US-0055108P.
XX
PA (CORR) CORNELL RES FOUND INC.
XX
PI Kim JF, Beer SV;
XX
DR WPI; 1999-167126/14.
XX
DR N-PSDB; AAX09007.
XX
PT New Erwinia amylovora hypersensitive response eliciting gene and protein
PT - useful for providing transgenic plants and seeds with enhanced growth,
PT and insect and disease resistance.
XX
PS Claim 1; Page 50-51; 54pp; English.
XX
CC The hypersensitive response eliciting protein (hrp) or polypeptide is
CC produced as part of an active defense by plants against incompatible
CC pathogen infections. The hypersensitive response is a rapid localised
CC necrosis. The hrp protein and gene when used in nucleotide constructs are
CC useful for providing disease resistance to plants, insect control to
CC plants, and enhancing plant growth (enhancing fruit size and earlier
CC colouration and maturation), by direct application of the protein to
CC plants, or by producing transgenic plants or seeds using the hrp gene
XX
SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57
RESULT 2
AAV71094
ID AAV71094 standard; protein; 447 AA.
AC AAV71094;
XX
DT 08-SEP-2000 (first entry)
XX
DE Erwinia amylovora hypersensitive response elicitor #2.

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57
RESULT 3
AAV84855
ID AAV84855 standard; protein; 447 AA.
AC AAV84855;
XX
DT 08-AUG-2000 (first entry)
XX
DE A hypersensitive response elicitor protein.
XX
KW Hypersensitive response; insect control; disease resistance;
KW hypersensitive response elicitor; plant growth; vegetable; crop;
KW ornamental plant.
XX
OS Erwinia amylovora.
XX
FN WO200020452-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US023181.
XX
PR 05-OCT-1998; 98US-0103050P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Fan H, Niggemeyer JL;
XX
DR WPI; 2000-303745/26.
XX
DR N-PSDB; AAA14939.

Query Match 100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57
RESULT 3
AAV84855
ID AAV84855 standard; protein; 447 AA.
AC AAV84855;
XX
DT 08-AUG-2000 (first entry)
XX
DE A hypersensitive response elicitor protein.
XX
KW Hypersensitive response; insect control; disease resistance;
KW hypersensitive response elicitor; plant growth; vegetable; crop;
KW ornamental plant.
XX
OS Erwinia amylovora.
XX
FN WO200020452-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US023181.
XX
PR 05-OCT-1998; 98US-0103050P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Fan H, Niggemeyer JL;
XX
DR WPI; 2000-303745/26.
XX
DR N-PSDB; AAA14939.

XX Hypersensitive response elicitor; environmental stress resistance; plant.
KW Erwinia amylovora.
OS WO200028055-A2.
XX
FN 18-MAY-2000.
XX
PD 04-NOV-1999; 99WO-US026039.
XX
PF 05-NOV-1998; 98US-0107243P.
XX
PR (EDEN-) EDEN BIOSCIENCE CORP.
XX
PA Wei Z, Schading RL;
XX
PI WPI; 2000-376566/32.
XX
DR N-PSDB; AAD00669.
XX
PT Application of a hypersensitive response elicitor protein to plants to
PT impart stress resistance.
XX
PS Disclosure; Page 10-12; 84pp; English.
XX
CC The patent discloses a method to impart stress resistance to plants by
CC applying a hypersensitive response elicitor in a non-infectious form to a
CC plant or seed. The present sequence is a hypersensitive response elicitor
CC protein from Erwinia amylovora. The protein is heat stable, protease
CC sensitive and suppressed by inhibitors of plant metabolism. It is used to
CC impart stress resistance to plants
XX
SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57
RESULT 3
AAV84855
ID AAV84855 standard; protein; 447 AA.
AC AAV84855;
XX
DT 08-AUG-2000 (first entry)
XX
DE A hypersensitive response elicitor protein.
XX
KW Hypersensitive response; insect control; disease resistance;
KW hypersensitive response elicitor; plant growth; vegetable; crop;
KW ornamental plant.
XX
OS Erwinia amylovora.
XX
FN WO200020452-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US023181.
XX
PR 05-OCT-1998; 98US-0103050P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Fan H, Niggemeyer JL;
XX
DR WPI; 2000-303745/26.
XX
DR N-PSDB; AAA14939.

Query Match 100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57
RESULT 3
AAV84855
ID AAV84855 standard; protein; 447 AA.
AC AAV84855;
XX
DT 08-AUG-2000 (first entry)
XX
DE A hypersensitive response elicitor protein.
XX
KW Hypersensitive response; insect control; disease resistance;
KW hypersensitive response elicitor; plant growth; vegetable; crop;
KW ornamental plant.
XX
OS Erwinia amylovora.
XX
FN WO200020452-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US023181.
XX
PR 05-OCT-1998; 98US-0103050P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Fan H, Niggemeyer JL;
XX
DR WPI; 2000-303745/26.
XX
DR N-PSDB; AAA14939.

Query Match 100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NSALGQQPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQQPIDRQTIEQMAQLLAELLS 57
RESULT 3
AAV84855
ID AAV84855 standard; protein; 447 AA.
AC AAV84855;
XX
DT 08-AUG-2000 (first entry)
XX
DE A hypersensitive response elicitor protein.
XX
KW Hypersensitive response; insect control; disease resistance;
KW hypersensitive response elicitor; plant growth; vegetable; crop;
KW ornamental plant.
XX
OS Erwinia amylovora.
XX
FN WO200020452-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-US023181.
XX
PR 05-OCT-1998; 98US-0103050P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Wei Z, Fan H, Niggemeyer JL;
XX
DR WPI; 2000-303745/26.
XX
DR N-PSDB; AAA14939.

XX Hypersensitive response elicitor polypeptides useful for imparting
PT enhanced growth, disease resistance and insect resistance to plants,
PT especially vegetables and ornamental flowers.
XX
XX PS Disclosure; Page 12-13; 100pp; English.
XX
XX CC The present sequence represents a hypersensitive response elicitor
CC polypeptide. The specification describes hypersensitive response elicitor
CC polypeptide fragments, which do not elicit a hypersensitive response.
CC Instead, the proteins impart disease resistance to plants, enhance plant
CC growth, and/or control insects. The polypeptide fragments may be used to
CC these properties to plants. The plants which may be treated in this way
CC include vegetables, crops and ornamental plants such as alfalfa, rice,
CC wheat, barley, rye, cotton, sunflower, peanut, corn, potato, sweet
CC potato, bean, pea, chickory, lettuce, endive, cabbage, brussel sprout,
CC beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion,
CC garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini,
CC cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry,
CC pineapple, soybean, tobacco, tomato, sorghum or sugarcane, Arabidopsis
CC thaliana, Saintpaulia, petunia, pelargonium, poinsettia, chrysanthemum,
CC carnation or zinnia
XX
XX SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 3; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQPPIDRQTIEQMAQLLAELLS 57

RESULT 4
AA022548
ID AAO22548 standard; protein; 447 AA.
XX
XX AC AAO22548;
XX
XX DT 28-OCT-2002 (first entry)
XX
XX DE Hypersensitive response elicitor protein #2 from Erwinia amylovora.
XX
XX KW Desiccation inhibitor; longevity enhancer; desiccation inhibitor;
KW cutting; ornamental plant; hypersensitive response elicitor protein;
KW transgenic ornamental plant; early flowering; longevity enhancer; plant;
KW flower bloom; flower.
XX
XX OS Erwinia amylovora.
XX
XX PN WO200237960-A2.
XX
XX PD 16-MAY-2002.
XX
XX PF 06-NOV-2001; 2001WO-US043715.
XX
XX PR 13-NOV-2000; 2000US-0248169P.
XX
XX PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
XX PI Wei Z, Leon E, Oviedo A;
XX
XX WPI; 2002-575194/61.
XX
XX DR N-PSDB; AAL41133.
XX
XX PT Inhibiting desiccation of cuttings from ornamental plants, by treating
PT ornamental plants with hypersensitive response elicitor protein, or
PT expressing heterologous hypersensitive response elicitor protein in
PT plants.
XX
XX PS Disclosure; Page 12-13; 69pp; English.
XX

CC The invention relates to a method for inhibiting desiccation of cuttings
CC from ornamental plants. The method involves treating the cuttings with a
CC hypersensitive response elicitor protein or polypeptide, or providing a
CC transgenic ornamental plant or plant seed transformed with a DNA molecule
CC encoding the hypersensitive response elicitor polypeptide, and growing
CC the ornamental plant or transgenic ornamental plant produced from the
CC transgenic ornamental plant seed. The hypersensitive response elicitor
CC protein or polypeptide is useful for inhibiting desiccation of cuttings
CC from ornamental plants, for harvesting cuttings from ornamental plants,
CC for promoting early flowering of ornamental plants, and enhancing the
CC longevity of flower blooms on ornamental plant cuttings. This sequence
CC represents a hypersensitive response elicitor protein of the invention
XX
XX SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;
Best Local Similarity 100.0%; Pred. No. 6.4e-11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLS 27
DB 31 NSALGQPPIDRQTIEQMAQLLAELLS 57

RESULT 5
AAE18296
ID AAE18296 standard; protein; 447 AA.
XX
XX AC AAE18296;
XX
XX DT 07-MAY-2002 (first entry)
XX
XX DE Erwinia amylovora hypersensitive response elicitor (HRE) #2.
XX
XX KW Hypersensitive response elicitor; HRE; transgenic plant; plant growth;
KW stress tolerance; disease tolerance; modified flower colour;
KW insect resistance; herbicide resistance; male sterility.
XX
XX OS Erwinia amylovora.
XX
XX PN WO200195724-A2.
XX
XX PD 20-DEC-2001.
XX
XX PF 13-JUN-2001; 2001WO-US018955.
XX
XX PR 15-JUN-2000; 2000US-0211585P.
XX
XX PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
XX PI Wei Z, Derocher J;
XX
XX WPI; 2002-130707/17.
XX
XX DR N-PSDB; AAD29125.
XX
XX PT Improving effectiveness of transgenic plants by topical application of a
PT hypersensitive response elicitor protein to the transgenic plant or by
PT incorporating into the plant a transgene encoding the protein.
XX
XX PS Disclosure; Page 13-14; 86pp; English.
XX
XX CC The invention relates to methods of improving the effectiveness of
CC transgenic plants which involves either topical application of a
CC hypersensitive response elicitor (HRE) protein to the transgenic plant or
CC incorporating into the transgenic plant a transgene encoding HRE. HRE
CC sequence is used for improving the effectiveness of transgenic plants by
CC maximising the benefit of transgenic traits associated with a deleterious
CC effect on growth, stress tolerance, disease or insect resistance.
CC enhanced growth, herbicide resistance, male sterility, modified flower
CC colour and biochemically modified plant product in the transgenic plants
CC or overcoming the deleterious effects. The present sequence is Erwinia
CC amylovora HRE protein
XX

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;
 Best Local Similarity 100.0%; Pred. No. 6,4e-11;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQGPIDRQTIEQMAQLLAELLS 27
 |||||
 Db 31 NSALGQGPIDRQTIEQMAQLLAELLS 57

RESULT 6
 AAE16448
 ID AAE16448 standard; protein; 447 AA.
 XX
 AC AAE16448;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE E. amylovora hypersensitive response elicitor protein #2.
 XX
 KW Hypersensitive response elicitor protein; plant growth; fruit coloration;
 KW disease resistance; stress resistance; phytotoxin; insect infection;
 KW plant maturation.
 XX
 OS Erwinia amylovora.
 XX

Key Location/Qualifiers
 FT Domain 5..64
 FT /label= Hypersensitive_response_eliciting_domain_1
 FT Region 5..45
 FT /label= Acidic_unit
 FT Region 45..64
 FT /label= Alpha_helix
 FT Domain 103..146
 FT /label= Hypersensitive_response_eliciting_domain_2
 FT Region 103..131
 FT /label= Acidic_unit
 FT Region 131..146
 FT /label= Alpha_helix
 XX
 PN WO200198501-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 12-JUN-2001; 2001WO-US018820.
 XX
 PR 16-JUN-2000; 2000US-0212211P.
 XX
 PA (EDEN-) EDEN BIOSCIENCE CORP.
 XX
 PI Fan H, Wei Z;
 XX
 WI 2002-122282/16.
 DR N-PSDB; AAD27016.
 XX

New hypersensitive response elicitor proteins comprising spaced apart domains having an acidic portion linked to an alpha-helix, useful for imparting disease or stress resistance, controlling insects or enhancing plant growth.

Disclosure; Page 13-14; 99pp; English.

The patent discloses hypersensitive response elicitor proteins and nucleotides encoding such proteins. Hypersensitive response elicitor proteins comprise an isolated pair or more of spaced apart domains, each comprising an acidic portion linked to an alpha-helix and capable of eliciting a hypersensitive response in plants. Sequences of the invention are used to impart disease resistance to plants, to enhance plant growth, to control insects and/or to impart stress resistance to plants which includes resistance to environmental stresses such as climate, air pollution, chemical and nutritional stress. The method of imparting disease resistance has the potential for treating previously untreatable

diseases, treating diseases systemically and avoiding the use of infectious agents or environmentally harmful materials. Hyper-sensitive response elicitor sequences are used to enhance plant growth which encompasses greater yield, increased in quantity of seeds produced, percentage of seeds germinated, plant size and biomass, bigger fruits, earlier fruit coloration and plant maturation. They are also used for insect control which encompasses preventing direct insect damage to plant by feeding injury, interfering with insect larval feeding on the plants, preventing insects from colonizing host plants and releasing phytotoxins. Sequences of the invention also prevent subsequent disease damage to plants resulting from insect infection. The present sequence is Erwinia amylovora hypersensitive response elicitor protein

SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;
 Best Local Similarity 100.0%; Pred. No. 6,4e-11;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQGPIDRQTIEQMAQLLAELLS 27
 |||||
 Db 31 NSALGQGPIDRQTIEQMAQLLAELLS 57

RESULT 7
 ABB09226
 ID ABB09226 standard; protein; 447 AA.
 XX
 AC ABB09226;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 DE Erwinia amylovora hypersensitive response elicitor SEQ ID NO:5.
 XX
 KW Hypersensitive response elicitor; fruit; vegetable; plant; desiccation;
 KW postharvest disease.
 XX
 OS Erwinia amylovora.
 XX
 PN WO200180639-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 17-APR-2001; 2001WO-US012468.
 XX
 PR 19-APR-2000; 2000US-0198359P.
 XX
 PA (EDEN-) EDEN BIOSCIENCE CORP.
 XX
 PI Wei Z, Qiu D, Remick D;
 XX
 WI 2002-041357/05.
 DR N-PSDB; ABL51711.
 XX

Inhibiting post harvest disease (caused by Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a fruits or vegetables, using hypersensitive response elicitor proteins or nucleic acids.

Disclosure; Page 11-12; 72pp; English.

The present invention describes methods for inhibiting post harvest disease or desiccation and enhancing the longevity in a fruits or vegetables, using hypersensitive response elicitor proteins or polypeptides or nucleic acids (i) derived from pathogens (e.g. Erwinia amylovora, E. stewartii, E. chrysanthemi, E. carotovora, Xanthomonas, Pseudomonas syringae, P. solanacearum, Phytophthora, and Clavibacter). (i) has bactericidal activity, and can be used in gene therapy. The method can be used for inhibiting post harvest disease (caused by Penicillium, Botrytis, Phytophthora, or Erwinia) or desiccation and enhancing the longevity in a fruits or vegetables. The method enables growers, warehouse packers, shippers and suppliers to process, handle and store fruit and vegetables with reduced losses caused by post harvest

CC disease and desiccation, therefore reducing costs to the consumer and
 CC improving quality. The present sequence represents a hyperresponsive
 CC response elicitor protein given in the exemplification of the present
 CC invention
 XX
 SQ Sequence 447 AA;

Query Match 100.0%; Score 128; DB 5; Length 447;
 Best Local Similarity 100.0%; Pred. No. 6.4e-11;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSALGQOPIDRTIQEQMAQLLAELLS 27
 |||||
 Db 31 NSALGQOPIDRTIQEQMAQLLAELLS 57

RESULT 8
 ABO73754
 ID ABO73754 standard; protein; 313 AA.
 XX
 AC ABO73754;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #5929.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX
 WPI; 2003-615309/58.
 DR N-PSDB; ABD07325.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 22500; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biotech technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 313 AA;

Query Match 40.6%; Score 52; DB 7; Length 313;
 Best Local Similarity 55.6%; Pred. No. 19;

Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 Qy 7 QPIDRTIQEQMAQLLAEL 24
 |||||
 Db 21 QRLERQVAVQAQVLAEL 38

RESULT 9
 ABO73936
 ID ABO73936 standard; protein; 457 AA.
 XX
 AC ABO73936;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #6111.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX
 WPI; 2003-615309/58.
 DR N-PSDB; ABD07507.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 22682; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biotech technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 457 AA;

Query Match 40.6%; Score 52; DB 7; Length 457;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 4 LGQOPIDRTIQEQMAQLLAEL 25
 |||||
 Db 312 LGQSPILALABELLAEL 333

RESULT 10
 ADN73161

CC above polypeptides a vector comprising (5); (6) a recombinant host cell
 CC containing the vector; and (10) a protein chip comprising Shigella
 CC flexneri polypeptide and a mammalian polypeptide defined in the
 CC specification. A pharmaceutical composition comprising the compound,
 CC polypeptide or polynucleotide is useful for treating or preventing
 CC shigellosis (bacillary dysentery) in a human or mammal. The present
 CC sequence represents a human prey protein isolated by the yeast two-hybrid
 CC assay, forming a complex of the invention with a shigella protein
 XX
 SQ Sequence 266 AA;

Query Match 39.5%; Score 50.5; DB 5; Length 266;
 Best Local Similarity 60.0%; Pred. No. 27;
 Matches 12; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 NSALGO-QPIDRQTIEQMAQ 19
 ||||| : : : : :
 Db 60 NSALGQTQPSDQDTLVQRAE 79

RESULT 19
 AAW22307
 ID AAW22307 standard; protein; 352 AA.
 XX
 AC AAW22307;
 XX
 DT 27-AUG-2003 (revised)
 DT 18-MAR-1998 (first entry)
 XX
 DE Lipase modulator from P. wisconsinensis.
 XX
 KW Pseudomonas wisconsinensis; lipase modulator; expression system;
 KW enzyme production; detergent.
 XX
 OS Pseudomonas sp.
 XX
 PN WO9713847-A1.
 XX
 PD 17-APR-1997.
 XX
 PF 14-OCT-1996; 96WO-BE000109.
 XX
 PR 12-OCT-1995; 95BE-00000851.
 XX
 PA (GENV) GENENCOR INT INC.
 XX

PI Charmoille L, Andre C, Hazbon MH, Cornelis P, Dhaese P;
 XX WPI; 1997-235883/21.
 DR N-PSDB; AAT73345.
 XX
 PT Expression system containing regulatory sequences from P. wisconsinensis
 PT - and optionally modulator and GPW protein sequences, especially for
 PT lipase production for use e.g. in detergents.
 XX
 PS Claim 23; Page 36-38; 74pp; French.
 XX
 CC This sequence, isolated from Pseudomonas wisconsinensis, is a lipase
 CC modulator. A claimed expression system for production of an enzyme,
 CC especially a Pseudomonas lipase, comprises at least sequences for a
 CC promoter, signal sequence, mature enzyme and a terminator. The
 CC transformed cells are used to produce enzymes, especially lipase for use
 CC in detergents. Other enzymes are useful in food, pharmaceutical and
 CC chemical industries. The system provides effective secretion of large
 CC quantities of enzyme into the culture medium. GPW seems to assist
 CC expression of enzyme and to protect it (and cells) against oxygen
 CC radicals, also to repair cell membranes damaged by oxygen radicals.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX

SQ Sequence 352 AA;
 Query Match 39.5%; Score 50.5; DB 2; Length 352;
 Best Local Similarity 61.9%; Pred. No. 38;

Matches 13; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 3 ALGOQPID-RQTIEQMAQLIA 22
 ||||| : : : : :
 Db 149 ALMOQPLDARQQVEPQALQA 169

RESULT 20
 ABG61825
 ID ABG61825 standard; protein; 596 AA.
 XX
 AC ABG61825;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE Prostate cancer-associated protein #26.
 XX
 KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
 XX
 OS Mammalia.
 XX
 PN WO200230268-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-US032045.
 XX
 PR 13-OCT-2000; 2000US-00687576.
 PR 08-DEC-2000; 2000US-00733288.
 PR 08-DEC-2000; 2000US-00733742.
 PR 24-JAN-2001; 2001US-0263957P.
 PR 16-MAR-2001; 2001US-0276791P.
 PR 16-MAR-2001; 2001US-0276888P.
 PR 06-APR-2001; 2001US-0281922P.
 PR 24-APR-2001; 2001US-0286214P.
 PR 30-APR-2001; 2001US-00847046.
 PR 04-MAY-2001; 2001US-0288589P.
 XX

(EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 XX WPI; 2002-471335/50.
 DR N-PSDB; ABK92140.
 XX
 PT Detecting a prostate cancer-associated transcript in a cell in a patient,
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 PT by determining if prostate cancer-associated genes are expressed in a
 PT prostate tissue.
 XX
 PS Claim 27; Page 322; 436pp; English.
 XX
 CC The present invention relates to methods of detecting a prostate cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with prostate cancer-
 CC associated polynucleotides (designated PC genes) that selectively
 CC hybridise to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer-
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins
 XX

SQ Sequence 596 AA;
 Query Match 39.5%; Score 50.5; DB 5; Length 596;
 Best Local Similarity 60.0%; Pred. No. 69;
 Matches 12; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 1 NSALGQ-OPIDRQTIEQMAQ 19
| | | | | : | :
Db 390 NSALGQTQPSDQDTLVQRAE 409

Search completed: January 25, 2005, 12:11:58
Job time : 68.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:08:41 ; Search time 30.6346 Seconds
(without alignments)
58.450 Million cell updates/sec

Title: US-09-879-248-6_COPY_31_57

Perfect score: 128
Sequence: 1 NSALGQPDRTQIEQMAQLLALLKS 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	3	US-09-120-927-2
2	128	100.0	447	4	US-09-431-614-6
3	52	40.6	313	4	US-09-252-991A-22500
4	52	40.6	457	4	US-09-252-991A-22682
5	51	39.8	1312	2	US-08-592-126-148
6	51	39.8	1312	2	US-08-687-080-51
7	51	39.8	1312	4	US-09-168-595-148
8	49	38.3	173	4	US-09-270-767-40996
9	49	38.3	173	4	US-09-270-767-56212
10	49	38.3	264	4	US-09-252-991A-20342
11	48.5	37.9	237	4	US-09-252-991A-19010
12	48.5	37.9	891	4	US-09-252-991A-25595
13	47	36.7	112	4	US-09-489-039A-9733
14	47	36.7	505	4	US-09-252-991A-23615
15	45	35.2	620	3	US-08-982-785A-10
16	45	35.2	620	4	US-09-629-498-10
17	45	35.2	631	4	US-09-252-991A-20418
18	45	35.2	686	3	US-08-982-785A-8
19	45	35.2	686	4	US-09-629-498-8
20	45	35.2	756	3	US-08-982-785A-2
21	45	35.2	756	4	US-09-629-498-2
22	45	35.2	1084	3	US-09-227-725A-3
23	45	35.2	1225	4	US-09-501-171-4
24	45	35.2	1247	4	US-09-501-171-6
25	44.5	34.8	122	4	US-09-248-796A-28164
26	44	34.4	130	4	US-09-107-532A-4295
27	44	34.4	255	4	US-09-252-991A-31162

319	34.4	1	US-08-125-746-1	Sequence 1, Appli
319	34.4	1	US-08-948-276-2	Sequence 2, Appli
320	34.4	1	US-08-125-746-3	Sequence 3, Appli
320	34.4	4	US-08-948-276-1	Sequence 1, Appli
320	34.4	6	5225537-4	Patent No. 5225537
325	34.4	3	US-09-248-796A-20030	Sequence 20030, A
325	34.4	3	US-09-324-096A-2	Sequence 2, Appli
327	34.4	3	US-09-324-096A-4	Sequence 4, Appli
327	34.4	3	US-09-324-096A-6	Sequence 6, Appli
328	34.4	4	US-09-248-796A-24205	Sequence 24205, A
378	34.4	3	US-09-710-279-1428	Sequence 1428, Ap
387	34.4	3	US-09-134-001C-3248	Sequence 3248, Ap
457	34.4	4	US-08-924-629C-6	Sequence 6, Appli
502	34.4	4	US-10-101-464A-945	Sequence 945, App
635	34.4	4	US-09-252-991A-20298	Sequence 20298, A
93	33.6	4	US-09-489-039A-7880	Sequence 7880, Ap
190	33.6	4	US-09-252-991A-29854	Sequence 29854, A
315	33.6	4	US-09-252-991A-31381	Sequence 31381, A
430	33.6	4	US-08-956-171B-5244	Sequence 5244, Ap
430	33.6	4	US-08-781-986A-5244	Sequence 5244, Ap
883	33.6	4	US-09-543-681A-6612	Sequence 6612, Ap
983	33.6	3	US-09-134-001C-3814	Sequence 3814, Ap
1282	33.6	4	US-09-543-681A-5419	Sequence 5419, Ap
3072	33.6	3	US-09-413-814-93	Sequence 93, Appl
3079	33.6	3	US-09-413-814-80	Sequence 80, Appl
350	33.2	4	US-09-248-796A-16002	Sequence 16002, A
36	32.8	2	US-08-806-203-1	Sequence 1, Appli
161	32.8	4	US-09-107-532A-6339	Sequence 6339, Ap
184	32.8	4	US-09-248-796A-15181	Sequence 15181, A
191	32.8	4	US-09-270-767-44048	Sequence 44048, A
235	32.8	4	US-09-270-767-33215	Sequence 33215, A
238	32.8	4	US-09-270-767-31899	Sequence 31899, A
254	32.8	4	US-09-270-767-47116	Sequence 47116, A
254	32.8	4	US-09-252-991A-26502	Sequence 26502, A
328	32.8	4	US-09-270-767-59817	Sequence 59817, A
397	32.8	4	US-09-252-991A-18356	Sequence 18356, A
593	32.8	1	US-08-208-036-17	Sequence 17, Appl
593	32.8	1	US-08-428-823-17	Sequence 17, Appl
876	32.8	2	US-08-633-476-2	Sequence 2, Appli
877	32.8	1	US-08-208-036-14	Sequence 14, Appl
877	32.8	1	US-08-428-823-14	Sequence 14, Appl
984	32.8	4	US-09-252-991A-17368	Sequence 17368, A
1021	32.8	4	US-10-101-464A-954	Sequence 954, App
1388	32.8	4	US-09-270-767-44387	Sequence 44387, A
1548	32.8	1	US-08-463-092B-7	Sequence 7, Appli
1548	32.8	2	US-08-460-907B-7	Sequence 7, Appli
2004	32.8	1	US-08-375-709-15	Sequence 15, Appl
2004	32.8	1	US-08-752-929-15	Sequence 15, Appl
2004	32.8	3	US-09-090-793-9	Sequence 9, Appli
2004	32.8	4	US-09-231-899-9	Sequence 9, Appli
227	32.4	4	US-08-414-625-4	Sequence 4, Appli
367	32.4	4	US-09-248-796A-19641	Sequence 19641, A
448	32.4	4	US-09-543-681A-7737	Sequence 7737, Ap
539	32.4	4	US-09-252-991A-21493	Sequence 21493, A
582	32.4	4	US-09-248-796A-19158	Sequence 19158, A
2568	32.4	4	US-09-866-108A-3	Sequence 3, Appli
92	32.0	4	US-09-543-681A-7692	Sequence 7692, Ap
104	32.0	4	US-09-270-767-32777	Sequence 32777, Ap
104	32.0	4	US-09-270-767-47994	Sequence 47994, A
185	32.0	4	US-09-583-110-4159	Sequence 4159, Ap
302	32.0	4	US-09-134-000C-5055	Sequence 5055, Ap
352	32.0	4	US-09-252-991A-17455	Sequence 17455, A
372	32.0	4	US-09-248-796A-14919	Sequence 14919, A
453	32.0	4	US-09-248-796A-14922	Sequence 14922, A
484	32.0	1	US-08-597-236-2	Sequence 2, Appli
484	32.0	1	US-08-746-682A-2	Sequence 2, Appli
692	32.0	4	US-09-252-991A-26724	Sequence 26724, A
720	32.0	4	US-09-394-272-14	Sequence 14, Appl
1253	32.0	4	US-09-489-039A-12097	Sequence 12097, A
1428	32.0	4	US-09-252-991A-30731	Sequence 30731, A
1618	32.0	1	US-07-853-913-4	Sequence 4, Appli
1618	32.0	1	US-09-538-092-1143	Sequence 1143, Ap
2442	32.0	3	US-09-514-247A-10	Sequence 10, Appl


```
QY 4 LQSQPIDRTIQMAQLLAELL 25
DB 312 LQSQPLLAALAEELLQLAEML 333

RESULT 5
US-08-592-126-148
; Sequence 148, Application US/08592126
; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: TRANS. OF RAD50 cDNA (SEQ. 54), NT.
; INDIVIDUAL ISOLATE: 389 TO 4324
; US-08-687-080-51

Query Match 39.8%; Score 51; DB 2; Length 1312;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDRTIQMAQLLAELLKS 27
DB 1237 LDRENTESLAHALVEIKS 1255

RESULT 7
US-09-168-595-148
; Sequence 148, Application US/09168595
; Patent No. 6555666
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/168,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,126
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615

QY 9 IDRTIQMAQLLAELLKS 27
DB 1237 LDRENTESLAHALVEIKS 1255

RESULT 6
US-08-687-080-51
; Sequence 51, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

Query Match 39.8%; Score 51; DB 2; Length 1312;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDRTIQMAQLLAELLKS 27
DB 1237 LDRENTESLAHALVEIKS 1255
```

; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
US-09-168-595-148

Query Match 39.8%; Score 51; DB 4; Length 1312;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 IDRTIEQMAQLLAELLS 27
|||:||||:|:|:
Db 1237 LDRENTSLAHALVEIKS 1255

RESULT 8
US-09-270-767-40996
; Sequence 40996, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40996
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-40996

Query Match 38.3%; Score 49; DB 4; Length 173;
Best Local Similarity 44.4%; Pred. No. 4.9;
Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 NSALGQPIDRTIEQMAQLLAELLS 27
|||:||||:|:|:
Db 127 NSMLNQPCSQQLLATQLLYARLLRS 153

RESULT 9
US-09-270-767-56212
; Sequence 56212, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56212
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-56212

Query Match 38.3%; Score 49; DB 4; Length 173;
Best Local Similarity 44.4%; Pred. No. 4.9;
Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 NSALGQPIDRTIEQMAQLLAELLS 27
|||:||||:|:|:
Db 127 NSMLNQPCSQQLLATQLLYARLLRS 153

RESULT 10
US-09-252-991A-20342
; Sequence 20342, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20342
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20342

Query Match 38.3%; Score 49; DB 4; Length 264;
Best Local Similarity 58.8%; Pred. No. 8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 ALGQQPFIDRTIEQMAQ 19
|||:||||:|:|:
Db 196 AAGQHPDHRETVEQEAQ 212

RESULT 11
US-09-252-991A-19010
; Sequence 19010, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19010
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19010

Query Match 37.9%; Score 48.5; DB 4; Length 237;
Best Local Similarity 46.2%; Pred. No. 8.4;
Matches 12; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

QY 2 SALGQQ-PIDRTIEQMAQLLAELIK 26
|||:||||:|:|:
Db 26 SALPQEKPLPQAFDDDPQLAELVLR 51

RESULT 12
US-09-252-991A-25595
; Sequence 25595, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2361S

Query Match          36.7%; Score 47; DB 4; Length 505;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 12; Conservative      8; Mismatches    4; Indels

Qy   3 ALGQQPIDRQTIEQM----AQLLAEELK 26
     |||:||||::|||:||||:|:|:|:|:|:|:|
Db   317 AVGPGRDORTVVQGADADAQVAFLE 344


RESULT 15
US-08-982-785A-10
; Sequence 10, Application US/08982785A
; Patent No. 6258929
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
; APPLICANT: Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
; TITLE OF INVENTION: NUCLEAR ACIDS AND DIAGNOSIS USING THEM
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,785A
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,556
; FILING DATE: 02-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 05311/018001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-982-785A-10

Query Match          35.2%; Score 45; DB 3; Length 620;
Best Local Similarity 43.5%; Pred. No. 92;
Matches 10; Conservative    3; Mismatches   10; Indels

Qy   2 SALGQQPDRQTIEQMQAQLLAEL 24
     |||||:||||::|||:||||:|:|:~|:|
Db   315 SALGSSEIDSKTVENCVCILRLNL 337


RESULT 16
US-09-629-498-10
; Sequence 10, Application US/09629498
; Patent No. 6797511
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
              Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
```

NUCLEIC ACIDS AND DIAGNOSIS USING THEM

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/629,498

FILING DATE: 31-Jul-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/982,785

FILING DATE: 02-DEC-1997

APPLICATION NUMBER: US 60/031,556

FILING DATE: 02-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/DOCKET NUMBER: 05311/018001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 620 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-09-629-498-10

Query Match 35.2%; Score 45; DB 4; Length 620;
 Best Local Similarity 43.5%; Pred. No. 92;
 Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 2 SALGQQPIDRQTIEQMAQLLAEL 24
 ||||| : : : : :
 Db 315 SALGSSSEIDSKTVENCVCILRLN 337

RESULT 17

US-09-252-991A-20418

Sequence 20418, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 20418

LENGTH: 631

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20418

Query Match 35.2%; Score 45; DB 4; Length 631;
 Best Local Similarity 43.5%; Pred. No. 94;
 Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 3 ALGQQPIDRQTIEQMAQLLAEL 25
 ||||| : : : : :
 Db 182 ALGEEHLGRTFLGQFLQALALL 204

RESULT 18

US-08-982-785A-8

Sequence 8, Application US/08982785A

Patent No. 6258929

GENERAL INFORMATION:

APPLICANT: Kosik, Kenneth S.

APPLICANT: Zhou, Jianhua

TITLE OF INVENTION: ALARM RELATED PEPTIDES AND

NUMBER OF INVENTION: NUCLEIC ACIDS AND DIAGNOSIS USING THEM

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/982,785A

FILING DATE: 02-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/031,556

FILING DATE: 02-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/DOCKET NUMBER: 05311/018001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 686 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-982-785A-8

Query Match 35.2%; Score 45; DB 3; Length 686;
 Best Local Similarity 43.5%; Pred. No. 1e+02;
 Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 2 SALGQQPIDRQTIEQMAQLLAEL 24
 ||||| : : : : :
 Db 312 SALGSSSEIDSKTVENCVCILRLN 334

RESULT 19

US-09-629-498-8

Sequence 8, Application US/09629498

Patent No. 6797511

GENERAL INFORMATION:

APPLICANT: Kosik, Kenneth S.

APPLICANT: Zhou, Jianhua

TITLE OF INVENTION: ALARM RELATED PEPTIDES AND

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:09:52 ; Search time 50.3654 Seconds
(without alignments)
193.681 Million cell updates/sec

Title: US-09-879-248-6_COPY_31_57
Perfect score: 128
Sequence: 1 NSALGQPPDRQTIEQMAQLLALLKS 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	9	US-09-835-684-5
2	128	100.0	447	9	US-09-880-371-5
3	128	100.0	447	9	US-09-879-248-6
4	128	100.0	447	14	US-10-010-390-5
5	128	100.0	447	15	US-10-441-736-6
6	128	100.0	447	17	US-10-847-142-5
7	51	39.8	660	15	US-10-406-686A-50
8	51	39.8	1312	14	US-10-393-602-148
9	51	39.8	1318	16	US-10-408-765A-666
10	50.5	39.5	266	14	US-10-043-487-245
11	50.5	39.5	596	14	US-10-205-823-222
12	49	38.3	628	16	US-10-437-963-174471
13	49	38.3	697	14	US-10-183-687-244
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 5, Appli
					Sequence 148, App
					Sequence 666, App
					Sequence 245, App
					Sequence 222, App
					Sequence 174471, App
					Sequence 244, App

14	49	38.3	3249	15	US-10-282-122A-67572	Sequence 67572, A
15	48.5	37.9	211	17	US-10-425-115-32898	Sequence 32898, A
16	48	37.5	96	17	US-10-425-115-267186	Sequence 267186, A
17	48	37.5	470	14	US-10-369-493-4813	Sequence 4813, Ap
18	48	37.5	478	14	US-10-369-493-7572	Sequence 7572, Ap
19	48	37.5	492	14	US-10-032-189-10	Sequence 10, Appl
20	48	37.5	495	14	US-10-032-189-68	Sequence 68, Appl
21	48	37.5	630	15	US-10-282-122A-49449	Sequence 49449, A
22	48	37.5	1419	16	US-10-473-115-2	Sequence 2, Appli
23	48	37.5	1419	16	US-10-408-765A-1810	Sequence 1810, Ap
24	48	37.5	2246	15	US-10-282-122A-69540	Sequence 69540, A
25	47	36.7	83	15	US-10-424-599-147641	Sequence 147641, A
26	46	35.9	31	16	US-10-745-069-52	Sequence 52, Appl
27	46	35.9	31	17	US-10-768-288A-52	Sequence 52, Appl
28	46	35.9	31	17	US-10-780-325A-52	Sequence 52, Appl
29	46	35.9	31	17	US-10-869-649-52	Sequence 52, Appl
30	46	35.9	32	16	US-10-745-069-51	Sequence 51, Appl
31	46	35.9	32	17	US-10-768-288A-51	Sequence 51, Appl
32	46	35.9	32	17	US-10-780-325A-51	Sequence 51, Appl
33	46	35.9	32	17	US-10-869-649-51	Sequence 51, Appl
34	46	35.9	33	16	US-10-745-069-50	Sequence 50, Appl
35	46	35.9	33	17	US-10-768-288A-50	Sequence 50, Appl
36	46	35.9	33	17	US-10-780-325A-50	Sequence 50, Appl
37	46	35.9	33	17	US-10-869-649-50	Sequence 50, Appl
38	46	35.9	34	16	US-10-745-069-49	Sequence 49, Appl
39	46	35.9	34	17	US-10-768-288A-49	Sequence 49, Appl
40	46	35.9	34	17	US-10-780-325A-49	Sequence 49, Appl
41	46	35.9	34	17	US-10-869-649-49	Sequence 49, Appl
42	46	35.9	36	16	US-10-745-069-47	Sequence 47, Appl
43	46	35.9	36	17	US-10-768-288A-47	Sequence 47, Appl
44	46	35.9	36	17	US-10-780-325A-47	Sequence 47, Appl
45	46	35.9	36	17	US-10-869-649-47	Sequence 47, Appl
46	46	35.9	38	15	US-10-424-599-262864	Sequence 262864, A
47	46	35.9	238	15	US-10-425-114-49572	Sequence 49572, A
48	46	35.9	259	15	US-10-425-114-49036	Sequence 49036, A
49	46	35.9	284	16	US-10-437-963-192612	Sequence 192612, A
50	46	35.9	348	14	US-10-156-761-9234	Sequence 9234, Ap
51	46	35.9	449	9	US-09-738-626-6150	Sequence 6150, Ap
52	46	35.9	482	15	US-10-425-114-43041	Sequence 43041, A
53	46	35.9	497	14	US-10-369-493-9870	Sequence 9870, Ap
54	46	35.9	566	15	US-10-425-114-56072	Sequence 56072, A
55	46	35.9	863	15	US-10-425-114-55432	Sequence 55432, A
56	46	35.9	1010	15	US-10-424-599-195812	Sequence 195812, A
57	46	35.9	1013	15	US-10-424-599-277819	Sequence 277819, A
58	45.5	35.5	249	15	US-10-282-122A-65161	Sequence 65161, A
59	45.5	35.5	249	15	US-10-282-122A-65658	Sequence 65658, A
60	45.5	35.5	351	14	US-10-156-761-8050	Sequence 8050, Ap
61	45.5	35.5	422	15	US-10-282-122A-51218	Sequence 51218, A
62	45.5	35.5	620	15	US-10-425-114-59476	Sequence 59476, A
63	45	35.2	29	16	US-10-745-069-54	Sequence 54, Appl
64	45	35.2	29	17	US-10-768-288A-54	Sequence 54, Appl
65	45	35.2	29	17	US-10-780-325A-54	Sequence 54, Appl
66	45	35.2	29	17	US-10-869-649-54	Sequence 54, Appl
67	45	35.2	30	16	US-10-745-069-53	Sequence 53, Appl
68	45	35.2	30	17	US-10-768-288A-53	Sequence 53, Appl
69	45	35.2	30	17	US-10-780-325A-53	Sequence 53, Appl
70	45	35.2	30	17	US-10-869-649-53	Sequence 53, Appl
71	45	35.2	36	16	US-10-745-069-85	Sequence 85, Appl
72	45	35.2	36	17	US-10-768-288A-85	Sequence 85, Appl
73	45	35.2	36	17	US-10-780-325A-85	Sequence 85, Appl
74	45	35.2	36	17	US-10-869-649-85	Sequence 85, Appl
75	45	35.2	50	17	US-10-776-013-614	Sequence 614, App
76	45	35.2	50	17	US-10-776-013-614	Sequence 2, Appli
77	45	35.2	95	9	US-09-757-713-2	Sequence 173202, A
78	45	35.2	128	15	US-10-424-599-173202	Sequence 282746, A
79	45	35.2	134	17	US-10-425-115-282746	Sequence 41859, A
80	45	35.2	136	15	US-10-425-114-41859	Sequence 45859, A
81	45	35.2	136	15	US-10-425-114-45859	Sequence 40481, A
82	45	35.2	147	16	US-10-767-701-40481	Sequence 171006, A
83	45	35.2	162	15	US-10-424-599-171006	Sequence 5093, Ap
84	45	35.2	252	9	US-09-738-626-5093	Sequence 4, Appli
85	45	35.2	307	14	US-10-190-471-4	Sequence 5850, Ap
86	45	35.2	489	9	US-09-815-242-5850	

87 45 35.2 561 17 US-10-425-115-282743 Sequence 282743,
88 45 35.2 705 10 US-09-907-907A-43 Sequence 43, Appl
89 45 35.2 872 14 US-10-369-493-8935 Sequence 8935, Ap
90 45 35.2 978 15 US-10-282-122A-44151 Sequence 44151, A
91 45 35.2 981 9 US-09-815-242-12211 Sequence 12211, A
92 45 35.2 1029 16 US-10-437-963-164029 Sequence 164029,
93 45 35.2 1084 13 US-10-071-900-3 Sequence 3, Appl
94 45 35.2 1818 15 US-10-282-122A-64245 Sequence 64245, A
95 45 35.2 2020 16 US-10-437-963-195615 Sequence 195615,
96 45 35.2 3390 14 US-10-369-493-21101 Sequence 21101, A
97 45 35.2 11300 16 US-10-250-304A-2 Sequence 2, Appl
98 44.5 34.8 739 17 US-10-425-115-211368 Sequence 211368,
99 44.5 34.8 1153 15 US-10-282-122A-61350 Sequence 61350, A
100 44 34.4 36 16 US-10-745-069-86 Sequence 86, Appl

ALIGNMENTS

RESULT 1
US-09-835-684-5
; Sequence 5, Application US/09835684
; Patent No. US20020019337A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/71
; CURRENT APPLICATION NUMBER: US/09/835,684
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-835-684-5

Query Match 100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57
|||||

RESULT 2
US-09-880-371-5
; Sequence 5, Application US/09880371
; Patent No. US20020059658A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Derocher, Jay
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 21829/91
; CURRENT APPLICATION NUMBER: US/09/880,371
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/211,585
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-880-371-5

Query Match 100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57
|||||

RESULT 3
US-09-879-248-6
; Sequence 6, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 21829/81
; CURRENT APPLICATION NUMBER: US/09/879,248
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/212,211
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-09-879-248-6

Query Match 100.0%; Score 128; DB 9; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57
|||||

RESULT 4
US-10-010-390-5
; Sequence 5, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/248,169
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-010-390-5

Query Match 100.0%; Score 128; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLKS 27
|||||

Db 31 NSALGQPPIDRQTIEQMAQLLAELLKS 57
|||||

```
RESULT 5
US-10-441-736-6
; Sequence 6, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Schading, Richard L.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS
; TITLE OF INVENTION: RESISTANCE
; FILE REFERENCE: 21829/203 (BBC-003)
; CURRENT APPLICATION NUMBER: US/10/441,736
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: 60/107,243
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 09/431,614
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-441-736-6

Query Match      100.0%; Score 128; DB 15; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQGPIDRQTIEQMAQLLAELLKS 27
Db 31 NSALGQGPIDRQTIEQMAQLLAELLKS 57

RESULT 6
US-10-847-142-5
; Sequence 5, Application US/10847142
; Publication No. US20040265442A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Qiu, Dewen
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/197
; CURRENT APPLICATION NUMBER: US/10/847,142
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/835,684
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Erwinia amylovora
US-10-847-142-5

Query Match      100.0%; Score 128; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQGPIDRQTIEQMAQLLAELLKS 27
Db 31 NSALGQGPIDRQTIEQMAQLLAELLKS 57

RESULT 7
US-10-406-686A-50
; Sequence 50, Application US/10406686A
; Publication No. US20040033586A1
; GENERAL INFORMATION:
```

```
; APPLICANT: CROOKE, HELEN RACHEL
; APPLICANT: SHEA, JACQUELINE ELIZABETH
; APPLICANT: FELDMAN, ROBERT GRAHAM
; APPLICANT: COUTEBROZE, SYLVAIN GABRIEL
; APPLICANT: LEGROS, FRANCOIS-XAVIER
; TITLE OF INVENTION: ATTENUATED GRAM NEGATIVE BACTERIA
; FILE REFERENCE: 454313-3171.1
; CURRENT APPLICATION NUMBER: US/10/406,686A
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/370,282
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-406-686A-50

Query Match      39.8%; Score 51; DB 15; Length 660;
Best Local Similarity 55.0%; Pred. No. 83;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAELL 25
Db 178 QQPINKQKIEQLKQLQOELL 197

RESULT 8
US-10-393-602-148
; Sequence 148, Application US/10393602
; Publication No. US20030170714A1
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Behlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/393,602
; FILING DATE: 19-Mar-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 148:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
; SEQUENCE DESCRIPTION: SEQ ID NO: 148:
```

1 NSALGQ-QPIDRQTIEQMAQ 19

TITLE OF INVENTION: Rice Nucleic Acid Mol
 TITLE OF INVENTION: Plants and Uses There
 FILE REFERENCES: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 174471
 LENGTH: 628

```
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72408C.1.pep
US-10-437-963-174471

Query Match      38.3%; Score 49; DB 16; Length 628;
Best Local Similarity 43.3%; Pred. No. 1.6e+02;
Matches 13; Conservative 4; Mismatches 9; Indels 4; Gaps 1;

QY      1 NSALGQGPIDRQT-----IEQMAQLLAELLK 26
      | | | | | | | | | | | | | | | | | | | |
Db      415 NDAMDQPPSDCLTRVRLLEQVASTIAELVK 444

RESULT 13
US-10-183-687-244
; Sequence 244, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvelli, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BRL458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 244
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-183-687-244

Query Match      38.3%; Score 49; DB 14; Length 697;
Best Local Similarity 43.3%; Pred. No. 1.8e+02;
Matches 13; Conservative 4; Mismatches 9; Indels 4; Gaps 1;

QY      1 NSALGQGPIDRQT-----IEQMAQLLAELLK 26
      | | | | | | | | | | | | | | | | | | | |
Db      484 NDAMDQPPSDCLTRVRLLEQVASTIAELVK 513

RESULT 14
US-10-282-122A-67572
; Sequence 67572, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Hasebeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67572
; LENGTH: 3249
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (650)..(650)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1081)..(1081)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1088)..(1088)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1114)..(1114)
; OTHER INFORMATION: X=any amino acid
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-67572

Query Match      38.3%; Score 49; DB 15; Length 3249;
Best Local Similarity 61.1%; Pred. No. 1e+03;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      10 DROTIEQMAQLLAELLK 27
      | | | | | | | | | | | | | | | | | | | |
Db      2637 DROTIEQMAHWNLLQA 2654

RESULT 15
US-10-425-115-328898
; Sequence 328898, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
```

; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 328898
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(211)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_63023C.1.pap
US-10-425-115-328898

Query Match 37.9%; Score 48.5; DB 17; Length 211;
Best Local Similarity 39.4%; Pred. No. 52;
Matches 13; Conservative 4; Mismatches 3; Indels 13; Gaps 1;

QY 3 ALGOQPIDRQTEIQ-----MAQLLA 22
Db 156 ALGQMPIDRVGLEKEATALSRMGLRLTAQVLA 188

RESULT 16

US-10-425-115-267186
; Sequence 267186, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 267186
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_175277C.1.pap
US-10-425-115-267186

Query Match 37.5%; Score 48; DB 17; Length 96;
Best Local Similarity 47.4%; Pred. No. 25;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 ALGOQPIDRQTEIQMAQLL 21
Db 8 SLGEHPLDKSTILQVAHQL 26

RESULT 17

US-10-369-493-4813
; Sequence 4813, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4813

; LENGTH: 470
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-369-493-4813

Query Match 37.5%; Score 48; DB 14; Length 470;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 13; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 2 SALGOQPIDRQTEIQMAQLLAELLKS 27
Db 315 SAAGKYPV--QTETMAAICIEAEKS 338

RESULT 18

US-10-369-493-7572
; Sequence 7572, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7572
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Burkholderia cepacia

US-10-369-493-7572

Query Match 37.5%; Score 48; DB 14; Length 478;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 13; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 2 SALGOQPIDRQTEIQMAQLLAELLKS 27
Db 321 SAAGKYPV--QTETMAAICIEAEKS 344

RESULT 19

US-10-032-189-10
; Sequence 10, Application US/10032189
; Publication No. US20030170630A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P

; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard A
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li

; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Gorman, Linda
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Rieger, Daniel K
; APPLICANT: Edinger, Shlomit R

```

; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-228
; CURRENT APPLICATION NUMBER: US/10/032,189
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,495
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/258,171
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/269,940
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/274,192
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/277,826
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/279,840
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/282,981
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/283,656
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/309,247
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/311,754
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,331
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (492)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-10-032-189-10

Query Match      37.5%; Score 48; DB 14; Length 492;
Best Local Similarity 37.0%; Pred. No. 1.6e+02;
Matches 10; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY      1 NSALGOQPIDRQTIQMAQLLAELLS 27
DB      134 NFGMGKRSIEERILRGSFLLAELRKT 160

RESULT 20
US-10-032-189-68
; Sequence 68, Application US/10032189
; Publication No. US20030170630A1
; GENERAL INFORMATION:
; APPLICANT: Alabrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard A
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie J

```

```

; APPLICANT: Boldog, Ferenc L
; APPLICANT: Gorman, Linda
; APPLICANT: Gangolli, Bsha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Rieger, Daniel K
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-228
; CURRENT APPLICATION NUMBER: US/10/032,189
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,495
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/258,171
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/269,940
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/274,192
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/277,826
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/279,840
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/282,981
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/283,656
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/309,247
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/311,754
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,331
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (216)..(226)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-10-032-189-68

Query Match      37.5%; Score 48; DB 14; Length 495;
Best Local Similarity 37.0%; Pred. No. 1.7e+02;
Matches 10; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY      1 NSALGOQPIDRQTIQMAQLLAELLS 27
DB      134 NFGMGKRSIEERILRGSFLLAELRKT 160

Search completed: January 25, 2005, 12:18:53
Job time : 53.3654 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:08:06 ; Search time 11.9423 Seconds
(without alignments)
217.534 Million cell updates/sec

Title: US-09-879-248-6_COPY_31_57

Perfect score: 128

Sequence: 1 NSALGQPIDRQTIEQMAQLLELLKS 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR 79:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	447	T18447	HrpW protein - Erw
2	51	39.8	580	G69656	2-succinyl-6-hydro
3	51	39.8	1312	T30845	probable DNA repair
4	50	39.1	430	AE1803	sugar binding prot
5	50	39.1	430	AF1429	sugar binding prot
6	50	39.1	2706	T28155	variant-specific s
7	49	38.3	257	S76537	hypothetical prote
8	49	38.3	902	AH2999	conserved hypothet
9	49	38.3	903	A96284	hypothetical prote
10	48.5	37.9	213	C82953	hypothetical prote
11	48.5	37.9	627	S76462	hypothetical prote
12	48.5	37.9	762	H83348	probable acylase p
13	48	37.5	350	S76212	hypothetical prote
14	47	36.7	246	D87487	uridylate kinase I
15	47	36.7	356	S01992	glutenin low molec
16	47	36.7	817	T01866	hypothetical prote
17	47	36.7	1122	T28130	hypothetical prote
18	46	35.9	36	PCBO	pancreatic hormone
19	46	35.9	59	PCSH	pancreatic hormone
20	46	35.9	75	G91013	hypothetical prote
21	46	35.9	75	A85858	hypothetical prote
22	46	35.9	75	B64988	hypothetical prote
23	46	35.9	159	AE2720	conserved hypothet
24	46	35.9	159	A97502	hypothetical prote
25	46	35.9	409	S76119	hypothetical prote
26	46	35.9	792	T29187	hypothetical prote
27	45.5	35.5	249	A81001	electron transfer
28	45.5	35.5	249	F82018	electron transfer
29	45.5	35.5	929	S75098	hypothetical prote

30	45	35.2	36	1	A61132	pancreatic hormone
31	45	35.2	36	1	C61132	pancreatic hormone
32	45	35.2	36	1	D61132	pancreatic hormone
33	45	35.2	66	1	PCCT	Mut7/nudix family
34	45	35.2	203	2	C95194	mutator protein li
35	45	35.2	203	2	H98060	conserved hypothet
36	45	35.2	455	2	B72339	hypothetical prote
37	45	35.2	510	2	AG2985	sugar ABC transpor
38	45	35.2	511	2	H98297	hypothetical prote
39	45	35.2	595	2	G96491	conserved hypothet
40	45	35.2	605	2	A82381	polyribonucleotide
41	45	35.2	705	2	S70691	conserved hypothet
42	45	35.2	978	2	B89971	neural plakophilin
43	45	35.2	1247	2	T42209	hypothetical prote
44	45	35.2	1818	1	S73852	COP9 signalosome c
45	44.5	34.8	397	2	F52302	probable portal pr
46	44.5	34.8	454	2	F90875	probable portal pr
47	44.5	34.8	833	2	H90821	unknown protein en
48	44.5	34.8	859	2	H85680	probable portal pr
49	44.5	34.8	859	2	H90909	probable portal pr
50	44.5	34.8	859	2	C90853	probable portal pr
51	44.5	34.8	861	2	B85844	hypothetical prote
52	44	34.4	58	2	S73287	hypothetical prote
53	44	34.4	201	2	B70616	hypothetical prote
54	44	34.4	209	2	A83524	probable hydrolase
55	44	34.4	216	2	D98240	probable hydrolase
56	44	34.4	226	2	F83307	probable hydrolase
57	44	34.4	229	2	B95287	probable hydrolase
58	44	34.4	229	2	AH3045	hydrolase (impor
59	44	34.4	262	2	S72841	hypothetical prote
60	44	34.4	262	2	F87071	conserved hypothet
61	44	34.4	299	2	A99497	heterodisulfide re
62	44	34.4	320	1	AQHUP	annexin V - bovine
63	44	34.4	320	1	S27214	probable oligopept
64	44	34.4	330	2	B71163	conserved hypothet
65	44	34.4	381	2	D89922	hypothetical prote
66	44	34.4	417	2	A72236	hypothetical prote
67	44	34.4	503	2	A82193	Sun/nucleolar prot
68	44	34.4	597	2	B69251	probable electron
69	44	34.4	789	2	S62172	SHE4 protein - yea
70	44	34.4	896	2	S76064	hypothetical prote
71	44	34.4	896	2	S59990	phycobilisome anch
72	44	34.4	1002	2	T46033	receptor protein k
73	44	34.4	1003	2	T05898	hypothetical prote
74	44	34.4	1058	2	T19282	hypothetical prote
75	44	34.4	1069	2	D81826	exodeoxyribonuclea
76	43.5	34.0	449	2	D75033	probable prolifera
77	43.5	34.0	1499	2	AC2555	hypothetical prote
78	43	33.6	36	1	PCFG	pancreatic hormone
79	43	33.6	36	2	B60413	pancreatic hormone
80	43	33.6	93	1	PCDG	pancreatic hormone
81	43	33.6	131	2	T49331	hypothetical prote
82	43	33.6	179	2	B69406	conserved hypothet
83	43	33.6	182	2	B83539	hypothetical prote
84	43	33.6	209	2	T34530	hypothetical prote
85	43	33.6	212	2	AG0854	probable sugar ald
86	43	33.6	324	2	AE3075	transcription regu
87	43	33.6	324	2	D98211	transcription regu
88	43	33.6	413	2	C89780	hypothetical prote
89	43	33.6	427	2	A12148	processing protein
90	43	33.6	437	2	E71290	probable flagellar
91	43	33.6	441	2	A40707	myosin protein - Ca
92	43	33.6	457	2	S52206	protein-export mem
93	43	33.6	507	2	D75029	hypothetical prote
94	43	33.6	507	2	S74641	sensor histidine k
95	43	33.6	556	2	C87609	secretogranin II -
96	43	33.6	617	2	S27389	excinnuclease ABC c
97	43	33.6	675	2	B52662	excinnuclease ABC c
98	43	33.6	675	2	F81095	excinnuclease ABC c
99	43	33.6	1128	2	A49960	bud emergence prot
100	43	33.6	1132	2	AD1809	phycobilisome core

ALIGNMENTS

RESULT 1
 T18447
 HrpW protein - Erwinia amylovora
 C:Species: Erwinia amylovora
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T18447
 R:Gaudriault, S.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z18936
 A:Accession: T18447
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-447 <GAU>
 A:Cross-references: UNIPROT:O54508; EMBL:Y13831; NID:e1293450; PID:e1293451; PIDN:CAA741
 A:Experimental source: strain CFBP1430, specific host Pommoideae
 C:Genetics:
 A>Note: hrpW
 Query Match 100.0%; Score 128; DB 2; Length 447;
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NSALGQGPIDRTIEQMAQLLAELLKS 27
 DB 31 NSALGQGPIDRTIEQMAQLLAELLKS 57
 RESULT 2
 G69656
 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (EC 4.1.3.-) [validated]
 N:Alternate names: menCF protein; menD protein
 N:Contains: 2-oxoglutarate decarboxylase (EC 4.1.1.71)
 C:Species: Bacillus subtilis
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Aug-2004
 C:Accession: G69656; A61649; S27509; T46640; T46641
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: G69656
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-580 <KUN>
 A:Cross-references: UNIPROT:P23970; GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB15060.
 A:Experimental source: strain 168
 R:Rowland, B.; Hill, K.; Mueller, J.; Driscoll, J.; Taber, H.
 Gene 167, 105-109, 1995
 A>Title: Structural organization of a Bacillus subtilis operon encoding menaquinone bios
 A:Reference number: A61649; MUID:96144257; PMID:8566759
 A:Accession: A61649
 A:Molecule type: DNA
 A:Residues: 1-111, 'R', '113-151', 'P', '153-539', 'RRQTS', '545', 'DSI', 'R02>
 A:Cross-references: GB:M74538; GB:M74182; GB:M74183; NID:g1185287; PIDN:ACAC37014.1; PID:
 A>Note: this is a revision to the sequence from reference S27507
 R:Rowland, B.; Hill, K.; Mueller, J.; Driscoll, J.; Taber, H.
 submitted to the EMBL Data Library, October 1991
 A:Description: Organization of an operon involved in menaquinone biosynthesis in Bacillu
 A:Reference number: S27507

A:Accession: S27509
 A:Molecule type: DNA
 A:Residues: 'MS', '30', 'LKIYT', '36-111', 'R', '113-164', 'ALSM', '<ROW1>
 A:Cross-references: EMBL:M74538; GB:M74182; GB:M74183; NID:g1185287
 A>Note: this sequence has been revised in reference A61649
 A:Accession: S27510
 A:Molecule type: DNA
 A:Residues: 178-539, 'RRQTS', '545', 'DSI', 'R0W2>
 A:Cross-references: EMBL:M74538; GB:M74182; GB:M74183; NID:g1185287
 A>Note: this sequence has been revised in reference A61649
 R:Driscoll, J.R.; Taber, H.W. 1992
 J. Bacteriol. 174, 5063-5071, 1992
 A>Title: Sequence organization and regulation of the Bacillus subtilis menBE operon.
 A:Reference number: A42715; MUID:92332443; PMID:1629163
 A:Accession: T46640
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'MS', '30', 'LKIYT', '36-111', 'R', '113-164', 'ALSM', '<DRI>
 A:Cross-references: EMBL:M74521; NID:g557486; PIDN:AAAS0398.1; PID:g557489
 A:Accession: T46641
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 178-539, 'RRQTS', '545', 'DSI', 'DR2>
 A:Cross-references: EMBL:M74521; NID:g557486; PIDN:AAAS0399.1; PID:g557490
 C:Genetics:
 A:Gene: menD; menCF
 A:Start codon: TTG
 C:Function:
 A:Pathway: ubiquinone biosynthesis
 A>Note: thiamine diphosphate cofactor
 C:Superfamily: 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase MenD
 C:Keywords: carbon-carbon lyase; carboxy-lyase; oxo-acid-lyase
 Query Match 39.8%; Score 51; DB 1; Length 580;
 Best Local Similarity 40.9%; Pred. No. 23;
 Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
 QY 5 GQPIDRTIEQMAQLLAELLK 26
 DB 200 GTQSVDRSLSDVAEMLAEEK 221
 RESULT 3
 T30845
 probable DNA repair protein RAD50 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T30845
 R:Kim, K.K.; Daud, A.I.; Wong, S.C.; Pajak, L.; Tsai, S.C.; Wang, H.; Henzel, W.J.; Fiecl
 J. Biol. Chem. 271, 29255-29264, 1996
 A>Title: Mouse RAD50 has limited epitopic homology to p53 and is expressed in the adult m
 A:Reference number: Z20899; MUID:97067183; PMID:8910585
 A:Accession: T30845
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1312 <KIM>
 A:Cross-references: UNIPROT:P70388; EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AACS281
 C:Genetics:
 A:Map position: 11
 A:Gene: RAD50
 C:Superfamily: RAD50 protein
 C:Keywords: DNA repair
 Query Match 39.8%; Score 51; DB 2; Length 1312;
 Best Local Similarity 47.4%; Pred. No. 55;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 9 IDRQTIEQMAQLLAELLKS 27
 DB 1237 LDRENIESLAHVEIINKS 1255
 RESULT 4

A:Reference number: Z20477; MUID:97373957; PMID:9230440
A:Accession: T28155
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2706 <ROW>
A:Cross-references: UNIPROT:O15870; EMBL:Y13402; PIDN:CAA73831.1
A:Experimental source: strain IT 4/25/5
C:Genetics:
A:Introns: 2493/3
A:Note: R29R+var1

Query Match 39.1%; Score 50; DB 2; Length 2706;
Best Local Similarity 47.6%; Pred. No. 1.6e+02;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 5 GQPIDRQTIEQMAQLLAELL 25
|:|:|:|:|:|:|:|:|:|:|:
Db 2108 GKTPIDAQNFQMGQTAKFV 2128

RESULT 7
S76537
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76537
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76537
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <KAN>
A:Cross-references: UNIPROT:Q55741; EMBL:D64002; GB:AB001339; NID:g1001612; PIDN:BAAL038
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 38.3%; Score 49; DB 2; Length 257;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 LGQOPIDRQTIEQMAQLLAELL 25
|:|:|:|:|:|:|:|:|:|:|:
Db 177 LGQQQKRSIQWAFLLGNLL 198

RESULT 8
AH2999
conserved hypothetical protein Atu3602 [imported] - *Agrobacterium tumefaciens* (strain C58)
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AH2999
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I.
stex, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH2999
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-902 <KUR>
A:Cross-references: UNIPROT:Q8U9X3; GB:AE008689; PIDN:AAL44414.1; PID:g17742013; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3602
A:Map position: linear chromosome

Query Match 37.5%; Score 48; DB 2; Length 350;
Best Local Similarity 40.7%; Pred. No. 36;
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 NSALGQOPIRQTEQMAQLLAELLS 27
DB 133 NSSIGQIQIERQSTSATPQASPELLPS 159

RESULT 14
D87487
uridylylate kinase [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
A/Accession: D87487
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: D87487
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-246 <STO>
A/Cross-references: UNIPROT:Q9A705; GB:AE005673; NID:gl3423374; PIDN:AAK23896.1; GSPDB:G
C/Genetics:
A/Gene: CCL921
C/Superfamily: uridylylate kinase

Query Match 36.7%; Score 47; DB 2; Length 246;
Best Local Similarity 38.5%; Pred. No. 34;
Matches 10; Conservative 8; Mismatches 6; Indels 2; Gaps 1;

QY 4 LGQOP--IDRQTIEQMAQLLAELLS 27
DB 23 MGDTPYGIDTNTVQSVAEDEVASIVKS 48

RESULT 15
S01992
glutinin low molecular weight chain precursor - wheat
C/Species: Triticum aestivum (common wheat)
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C/Accession: S01992
R/Pitta, E.G.; Rafaleki, J.A.; Hedgcoth, C.
Nucleic Acids Res. 15, 11376, 1988
A/Title: Nucleotide sequence and encoded amino acid sequence of a genomic gene region fo
A/Reference number: S01992; MUID:89083577; PMID:3205747
A/Accession: S01992
A/Molecule type: DNA
A/Residues: 1-356 <PIT>
A/Cross-references: UNIPROT:P10385; EMBL:X07747; NID:g21782; PIDN:CAA30570.1; PID:g21783
C/Superfamily: gliadin

Query Match 36.7%; Score 47; DB 2; Length 356;
Best Local Similarity 58.8%; Pred. No. 50;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 LGQOPIDRQTIEQMAQL 20
DB 301 LGQOPQQQLAHQIAQL 317

RESULT 16
T01866
hypothetical protein T24M8.3 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01866
R/Latrelle, P.; Elliott, G.; Le, T.
submitted to the EMBL Data Library, August 1998

A/Description: The sequence of A. thaliana T24M8.
A/Reference number: Z14449
A/Accession: T01866
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-817 <LAT>
A/Cross-references: UNIPROT:O81512; EMBL:AF077409; NID:g3319365; PID:g3319369
A/Experimental source: cultivar Columbia
C/Genetics:
A/Map position: 4
A/Introns: 117/3; 188/2; 269/3; 307/3; 357/3; 395/3; 422/3; 447/1; 486/3; 513/3; 541/3; 5
A/Note: T24M8.3

Query Match 36.7%; Score 47; DB 2; Length 817;
Best Local Similarity 44.4%; Pred. No. 1.2e+02;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 8 PIDRQTIEQMAQLLAEL 25
DB 411 PIDQSEEQAKIMSELI 428

RESULT 17
T28130
hypothetical protein ZK970.6 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T28130
R/Berks, M.
submitted to the EMBL Data Library, April 1995
A/Reference number: Z20473
A/Accession: T28130
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1122 <WIL>
A/Cross-references: UNIPROT:Q23682; EMBL:Z49073; PIDN:CAA88890.1; GSPDB:GNO0020; CESP:ZK9
A/Experimental source: clone ZK970
C/Genetics:
A/Gene: CESP:ZK970.6
A/Map position: 2
A/Introns: 28/3; 72/2; 153/2; 281/1; 312/3; 354/3; 401/1; 442/3; 660/3; 761/1; 819/2; 95
C/Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain homo

Query Match 36.7%; Score 47; DB 2; Length 1122;
Best Local Similarity 47.4%; Pred. No. 1.7e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAEL 24
DB 798 EQEDRPKAEQICKLUSEM 816

RESULT 18
PC80
pancreatic hormone - bovine
N/Alternate names: pancreatic polypeptide
C/Species: Bos primigenius taurus (cattle)
C/Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 10-May-2001
C/Accession: A01570
R/Chance, R.E.; Johnson, M.G.; Hoffmann, J.A.; Jones, W.E.; Koffenberger Jr., J.E.
unpublished results, cited by Chance, R.E., Moon, N.E., and Johnson, M.G., in Methods of
San Francisco, and London, 1979
A/Reference number: A94465
A/Accession: A01570
A/Molecule type: protein
A/Residues: 1-36 <CHA>
C/Superfamily: pancreatic hormone
C/Keywords: amidated carboxyl end; hormone; pancreas
F;1-36/Product: pancreatic hormone #status experimental <PCH>
F;36/Modified site: amidated carboxyl end (Tyr) #status experimental

Query Match 35.9%; Score 46; DB 1; Length 36;
Best Local Similarity 57.9%; Pred. No. 6;

Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAEL 24
: | | | | | | | | | |
Db 6 EYPGDNATPEQMAQYAAEL 24

RESULT 19
PCSH
pancreatic hormone precursor - sheep (tentative sequence) (fragment)
N;Alternate names: pancreatic polypeptide precursor
N;Contains: pancreatic hormone; pancreatic icosapeptide
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C;Accession: B94465; A91323; A01569
R;Chance, R.E.; Johnson, M.G.; Hoffmann, J.A.; Jones, W.E.; Koffenberger Jr., J.E.
unpublished results, cited by Chance, R.E., Moon, N.E., and Johnson, M.G., in Methods of
San Francisco, and London, 1979
A;Reference number: A94465
A;Accession: B94465
A;Molecule type: protein
A;Residues: 1-36 <CHA>
A;Cross-references: UNIPROT:P01301
R;Schwartz, T.W.; Hansen, H.F.
FEBS Lett. 168, 293-298, 1984
A;Title: Isolation of ovine pancreatic icosapeptide: a peptide product containing one cy
A;Reference number: A91323; MOID:84208777; PMID:6723953
A;Accession: A91323
A;Molecule type: protein
A;Residues: 40-59 <SCH>
C;Comment: We have added, by homology with the human precursor sequence, Gly-37, which i
C;Superfamily: pancreatic hormone
C;Keywords: amidated carboxyl end; hormone; pancreas
F:1-36/Product: pancreatic hormone #status experimental <PCH>
F:40-59/Product: pancreatic icosapeptide #status experimental <PCI>
F:36/Modified site: amidated carboxyl end (Tyr) (amide in mature form from following gly

Query Match 35.9%; Score 46; DB 1; Length 59;
Best Local Similarity 57.9%; Pred. No. 10;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 QQPIDRQTIEQMAQLLAEL 24
: | | | | | | | | | |
Db 6 EYPGDNATPEQMAQYAAEL 24

RESULT 20
G91013
hypothetical protein ECS3079 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: G91013
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: G91013
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-75 <HAV>
A;Cross-references: UNIPROT:P33921; GB:BA000007; PIDN:BA36502.1; PID:gl3362548; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECE3079
C;Superfamily: hypothetical protein HI0840

Query Match 35.9%; Score 46; DB 2; Length 75;
Best Local Similarity 64.7%; Pred. No. 13;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 IDRQTIEQMAQLLAEL 25
: | | : | | | | | | | |

Search completed: January 25, 2005, 12:15:55
Job time : 15.9423 secs

Db 4 ISRYSDQEQLLAEL 20

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2005, 12:07:26 ; Search time 101.769 Seconds
(without alignments)
152.650 Million cell updates/sec

Title: us-09-879-248-6_COPY_31_57

Perfect score: 128
Sequence: 1 NSALGQQPIDRQTIEQMAQLLAELLKS 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128	100.0	138	2	Q79AW7
2	128	100.0	447	2	O54508
3	128	100.0	447	2	Q9LAW2
4	109	85.2	450	2	O6XDB8
5	109	85.2	450	2	AAQ17046
6	109	85.2	450	2	AAS45453
7	61.5	48.0	1214	2	Q8D6Z1
8	53	41.4	416	2	O6DEJ7
9	53	41.4	440	2	O6IQ77
10	53	41.4	440	2	AAH71533
11	53	41.4	943	2	O7SRG5
12	52.5	41.0	452	2	Q8U310
13	51	39.8	377	2	O7SY40
14	51	39.8	413	2	O7NHF2
15	51	39.8	498	2	O8JHT9
16	51	39.8	580	1	MEND_BACSU
17	51	39.8	660	2	O9CKZ1
18	51	39.8	845	2	Q8BTF5
19	51	39.8	911	2	Q8BTF7
20	51	39.8	1017	2	O6VG40
21	51	39.8	1017	2	AAO23277
22	51	39.8	1173	2	O9UP86
23	51	39.8	1312	2	Q92878
24	51	39.8	1312	2	F70388
25	51	39.8	1312	2	Q9JTL8
26	51	39.8	1318	2	O43254
27	50.5	39.5	352	1	LIFO_PSEWI
28	50.5	39.5	596	1	ENH_HUMAN
29	50.5	39.5	596	2	AAT67739
30	50.5	39.5	1263	2	Q7UM20
31	50	39.1	430	2	Q926R9

32	50	39.1	430	2	O8Y3J6
33	50	39.1	430	2	Q71VS2
34	50	39.1	430	2	AAT05594
35	50	39.1	544	2	O97473
36	50	39.1	559	2	Q882U6
37	50	39.1	2706	2	O15870
38	49.5	38.7	477	2	O6RJC5
39	49	38.3	257	2	Q55741
40	49	38.3	305	2	Q88Y51
41	49	38.3	358	2	Q824U8
42	49	38.3	457	2	Q7X5J7
43	49	38.3	560	2	Q804Z3
44	49	38.3	628	2	Q84LP4
45	49	38.3	687	2	Q81MA8
46	49	38.3	902	2	Q8U9X3
47	49	38.3	903	2	Q7CSW1
48	49	38.3	1016	2	Q7R4T6
49	49	38.3	3470	2	Q88F79
50	48.5	37.9	213	2	Q9HT40
51	48.5	37.9	581	2	Q7VZV8
52	48.5	37.9	581	2	Q7WC19
53	48.5	37.9	581	2	Q7WQJ5
54	48.5	37.9	627	2	P74489
55	48.5	37.9	762	2	Q91194
56	48	37.5	223	2	Q9V6N0
57	48	37.5	302	2	Q6ZES1
58	48	37.5	302	2	BAD01829
59	48	37.5	307	2	Q88SC2
60	48	37.5	317	2	Q7R711
61	48	37.5	317	2	Q82KJ5
62	48	37.5	350	2	P74376
63	48	37.5	441	2	Q88LY0
64	48	37.5	463	2	Q8MMJ0
65	48	37.5	491	2	Q8WNE1
66	48	37.5	495	2	Q8WJ2
67	48	37.5	522	2	Q7WDI5
68	48	37.5	522	2	Q8D6M6
69	48	37.5	584	2	Q700Z9
70	48	37.5	584	2	CAD60270
71	48	37.5	1419	1	ESTA_HUMAN
72	48	37.5	4336	2	Q884F8
73	47.5	37.1	78	2	Q6T8R7
74	47.5	37.1	78	2	AAR11453
75	47.5	37.1	131	1	PAHO_BOVIN
76	47.5	37.1	187	2	Q8DJL3
77	47.5	37.1	604	2	Q74E16
78	47.5	37.1	604	2	AAR34524
79	47	36.7	148	2	Q721R5
80	47	36.7	148	2	AAS81409
81	47	36.7	202	2	Q8KW98
82	47	36.7	244	2	Q856T2
83	47	36.7	246	1	PYRH_CAUCR
84	47	36.7	281	2	Q871J0
85	47	36.7	348	2	Q95NK8
86	47	36.7	348	2	Q966U2
87	47	36.7	348	2	Q966U3
88	47	36.7	356	1	GLTA_WHEAT
89	47	36.7	357	2	Q6SPY8
90	47	36.7	357	2	AAS10192
91	47	36.7	437	2	Q6W3E3
92	47	36.7	437	2	AAQ82184
93	47	36.7	444	2	Q8XH51
94	47	36.7	476	2	Q82880
95	47	36.7	476	2	Q7BTF1
96	47	36.7	476	2	Q921P8
97	47	36.7	476	2	Q69097
98	47	36.7	487	2	Q9FNR5
99	47	36.7	502	2	Q56831
100	47	36.7	506	2	Q9KH29

ALIGNMENTS

```
RESULT 1
Q79AW7 ID Q79AW7 PRELIMINARY; PRT; 138 AA.
AC Q79AW7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW (Fragment).
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Ea321;
RX MEDLINE=98115919; PubMed=9448330;
RA Bogdanove A.J., Kim J.F., Wei Z., Kolchinsky P., Charkowski A.O.,
RA Conlin A.K., Collmer A., Beer S.V.;
RT "Homology and functional similarity of an hrp-linked pathogenicity
RT locus, dspEF, of Erwinia amylovora and the avirulence locus avrE of
RT Pseudomonas syringae pathovar tomato.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1325-1330(1998).
DR EMBL; U97504; AAC04849.1; -.
DR NCBI; 138
FT NON TER
SQ SEQUENCE 138 AA; 13788 MW; 4FE177177C74B3C6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLS 27
Db 31 NSALGQPPIDRQTIEQMAQLLAELLS 57

RESULT 2
O54508 ID O54508 PRELIMINARY; PRT; 447 AA.
AC O54508;
DT 01-NOV-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW protein.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CFBP1430;
RX MEDLINE=98086111; PubMed=9426142;
RA Gaudriault S., Malandrin L., Paulin J.P., Barny M.A.;
RT "DspA, an essential pathogenicity factor of Erwinia amylovora showing
RT homology with AvrE of Pseudomonas syringae, is secreted via the Hrp
RT secretion pathway in a DspB-dependent way.";
RL Mol. Microbiol. 26:1057-1069(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CFBP1430;
RX MEDLINE=98316710; PubMed=96541138;
RA Gaudriault S., Brisset M.N., Barny M.A.;
RT "HrpW of Erwinia amylovora, a new Hrp-secreted protein.";
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Ea321;
RX Kim J.F., Zumoff C.H., Beer S.V.;
RT "HrpW, a new harpin of Erwinia amylovora, is a member of a family of
RT pectate lyases.";
RL Phytopathology 87:0-0(1997).
```

```
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=Ea321;
RX MEDLINE=98422475; PubMed=9748455;
RA Kim J.F., Beer S.V.;
RT "HrpW of Erwinia amylovora, a new harpin that contains a domain
RT homologous to pectate lyases of a distinct class.";
RL J. Bacteriol. 180:5203-5210(1998).
DR EMBL; Y13831; CAA74158.1; -.
DR EMBL; U94513; AAC62314.1; -.
DR PIR; T18447; T18447.
DR HSSP; Q9RH00; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45347 MW; 7A744D169D6D9CF3 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLS 27
Db 31 NSALGQPPIDRQTIEQMAQLLAELLS 57

RESULT 3
Q9LAW2 ID Q9LAW2 PRELIMINARY; PRT; 447 AA.
AC Q9LAW2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Harpin HrpW.
GN Name=hrpW;
OS Erwinia amylovora.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=552;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Ea246;
RA Kim J.F., Laby R.J., Beer S.V.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083620; AAF63402.1; -.
DR HSSP; Q9RH00; 1EE6.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0030570; F:pectate lyase activity; IEA.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR004898; Pect_lyase.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 447 AA; 45340 MW; 0BBAAE3871EDC2F6 CRC64;

Query Match 100.0%; Score 128; DB 2; Length 447;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLS 27
Db 31 NSALGQPPIDRQTIEQMAQLLAELLS 57

RESULT 4
Q6XDB8 ID Q6XDB8 PRELIMINARY; PRT; 450 AA.
AC Q6XDB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HrpW.
GN Name=hrpW;
```

```
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
DR EMBL; AY530755; AAS45453.1; -.
DR InterPro; IPR011050; Pectin_lyase_like.
DR Pfam; PF03211; Pectate_lyase; 1.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match      85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLK 26
DB 31 NSAQGHPIRDTIEQMAQLLGELLK 56

RESULT 5
AAQ17046
ID AAQ17046 PRELIMINARY; PRT; 450 AA.
AC AAQ17046;
DT 03-MAR-2004 (T-EMBLrel. 27, Created)
DT 03-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DE HrpW.
GN Erwinia pyrifoliae.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Bae H.N., Cho J.M., Park D.H., Hwang I.K.,
RA Hur J.H., Lim C.K.;
RT "Molecular characterization hrp genes cluster of Erwinia pyrifoliae
RT and expression of hrpP encoding elicitor of the hypersensitive
RT response."
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237642; AAQ17046.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match      85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLK 26
DB 31 NSAQGHPIRDTIEQMAQLLGELLK 56

RESULT 6
AAS45453
ID AAS45453 PRELIMINARY; PRT; 450 AA.
AC AAS45453;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
```

```
DE HrpW.
GN HRPW.
OS Erwinia pyrifoliae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Erwinia.
OX NCBI_TaxID=79967;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WT#3;
RA Shrestha R., Baek S.J., Park D.H., Bae H.N., Cho J.M., Hwang I.,
RA Hur J.H., Lim C.K.;
RT "Identification of hrp genes cluster and characterization of HR
RT elicitor hrpNEP gene in Erwinia pyrifoliae."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY530755; AAS45453.1; -.
SQ SEQUENCE 450 AA; 45509 MW; 085F90CD322C3A74 CRC64;

Query Match      85.2%; Score 109; DB 2; Length 450;
Best Local Similarity 88.5%; Pred. No. 1.6e-06;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NSALGQPPIDRQTIEQMAQLLAELLK 26
DB 31 NSAQGHPIRDTIEQMAQLLGELLK 56

RESULT 7
Q8DGZ1
ID Q8DGZ1 PRELIMINARY; PRT; 1214 AA.
AC Q8DGZ1;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hydantoinease / oxoprolinase.
GN OrderedLocusNames=tlr2171;
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005376; BAC09723.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002821; Hydantoinease_A.
DR InterPro; IPR003692; Hydantoinease_B.
DR InterPro; IPR008040; Hydant_A_N.
DR Pfam; PF01968; Hydantoinease_A; 1.
DR Pfam; PF02538; Hydantoinease_B; 1.
DR Pfam; PF05378; Hydant_A_N; 1.
KW Complete proteome.
SQ SEQUENCE 1214 AA; 129918 MW; 929531C290F24447 CRC64;

Query Match      48.0%; Score 61.5; DB 2; Length 1214;
Best Local Similarity 54.2%; Pred. No. 15;
Matches 13; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 5 GQPIDRDTIE-QMAQLLAELLK 27
DB 395 GQPLDRDSVEQPAQLRQEIYRS 418

RESULT 8
Q6DEJ7
ID Q6DEJ7 PRELIMINARY; PRT; 416 AA.
AC Q6DEJ7;
```


DT	01-OCT-2004 (TReMBLrel. 28, Created)
DT	01-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT	01-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE	Hypothetical protein.
OS	Brachydanio rerio (Zebrafish) (Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Singapore local strain; TISSUE=Embryo;
RX	PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Haieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA	Rodriguez A.C., Grimwood J., Smutuz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Schmitt D.B., Schnerch A., Schein J.E.,
RA	Jones S.J., Matra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Singapore local strain; TISSUE=Embryo;
RA	Strausberg R.;
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC077115; AAH77115.1; -.
KW	Hypothetical protein.
QY	SEQUENCE 416 AA; 47726 MW; 008C38F5B606CE42 CRC64;
DB	Query Match 41.4%; Score 53; DB 2; Length 416;
DB	Best Local Similarity 38.5%; Pred. No. 79;
DB	Matches 10; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY	1 NSALGQQPDRQTQEQMAQLAELK 26
DB	1 NFGLGKKSIESRVLEESQYLIAELK 82
RESULT 9	
Q6IQ77	
ID	Q6IQ77 PRELIMINARY; PRT; 440 AA.
AC	Q6IQ77;
DT	05-JUL-2004 (TReMBLrel. 27, Created)
DT	05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TReMBLrel. 27, Last annotation update)
ZGC	86915.
GN	Name=zgc:86915;
OS	Brachydanio rerio (Zebrafish) (Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Embryo;
RX	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Haieh F.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071533; AAH71533.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 440 AA; 50312 MW; D1D4740B11E9FF193 CRC64;
 Query Match 41.4%; Score 53; DB 2; Length 440;
 Best Local Similarity 42.3%; Pred. No. 84;
 Matches 11; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
 QY 1 NSALGQOPIDRTQIEQMAQLLAELLK 26
 DB 81 NFGLGKKSIESRVLESQYLFALLK 106
 RESULT 11
 Q7SBG5 PRELIMINARY; PRT; 943 AA.
 ID Q7SBG5
 AC Q7SBG5;
 DT 01-MAR-2004 (T-EMBLrel. 26, Created)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Neurospora crassa.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
 RA Kamal M., Kamveselis M., Mauceli E., Bieleke C., Rudd S., Frishman D.,
 RA Kryzofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken D.,
 RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
 RL Nature 0:0-0(2003).
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; ABE01000168; EAA33735.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 943 AA; 105278 MW; 53586D8C6BDB7AF6 CRC64;
 Query Match 41.4%; Score 53; DB 2; Length 943;
 Best Local Similarity 40.9%; Pred. No. 1.8e+02;
 Matches 9; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 QY 4 LGQOPIDRTQIEQMAQLLAELL 25
 DB 536 VGRQPLDRVLQTLSQLREIL 557
 RESULT 12
 Q8U310 PRELIMINARY; PRT; 452 AA.
 ID Q8U310;
 AC Q8U310;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Putative nucleolar protein IV (Nol1-nop2-sun family).
 GN OrderedLocusNames=PF0666;
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 3638;
 RX MEDLINE=21079003; PubMed=11210495;
 RA Robb F.T., Maeder D.L., Brown J.R., Diruggiero J., Stump M.D.,
 RA Yeh R.K., Weiss R.B., Dunn D.M.;
 RT "Genomic sequence of hyperthermophile, Pyrococcus furiosus:
 RT implications for physiology and enzymology.";
 RL Meth. Enzymol. 330:134-157(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE010187; AAL80790.1; -;
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf.; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001678; Fmu_NOL1/Nop2p.
 DR InterPro; IPR011023; Nop2p.
 DR InterPro; IPR006027; NusB.
 DR InterPro; IPR006174; NusB region.
 DR InterPro; IPR000051; SAM_bind.
 DR Pfam; PF01189; NOL1_Nop2_Fmu; 1.
 DR ProDom; PD005242; NusB_region; 1.
 DR TIGRFAMs; TIGR00446; nop2p; 1.
 KW Complete proteome.
 SQ SEQUENCE 452 AA; 52212 MW; 99FFDD7419D4583E CRC64;
 Query Match 41.0%; Score 52.5; DB 2; Length 452;
 Best Local Similarity 40.8%; Pred. No. 1e+02;
 Matches 13; Conservative 6; Mismatches 8; Indels 5; Gaps 1;
 QY 1 NSALGQOPIDRTQIEQMAQLLAELLK 27
 DB 344 SGTIGKPELWRLRQKIEEWAQLQRELLS 375
 RESULT 13
 Q7SY40 PRELIMINARY; PRT; 377 AA.
 ID Q7SY40
 AC Q7SY40;
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Zgc:63546.
 GN Name=zgc:63546;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore J., Max S.I., Wang J., Heich F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,


```

Query Match      39.8%; Score 51; DB 2; Length 660;
Best Local Similarity 55.0%; Pred. No. 2.4e+02;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 6 QQPIDRQTIEOMAQLLAEL 25
Db 178 QQPHNKQIEQLKKLQOELL 197

RESULT 18
ID QBTF5 PRELIMINARY; PRT; 845 AA.
AC QBTF5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-
DE length enriched library, clone:C920014N24 product:DNA LIGASE IV (EC
DE 6.5.1.1) (POLYDEOXYRIBONUCLEOTIDE SYNTHASE).
GN Name-Lig4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

```

```

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Sano H.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + [deoxyribonucleotide] (N) +
CC [deoxyribonucleotide] (M) = AMP + diphosphate +
CC [deoxyribonucleotide] (N+M).
CC -!- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.
DR EMBL; AK083353; BAC38882.1; -.
DR HSP; P49917; I1K9.
DR MGD; MGI:1335098; Lig4.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR008994; Nucleic_acid_ob.
DR Pfam; PF00533; BRCT; 2.
DR Pfam; PF04679; DNA_ligase_A_C; 1.
DR Pfam; PF01068; DNA_ligase_A_M; 1.
DR Pfam; PF04675; DNA_ligase_A_N; 1.
DR SMART; SM00292; BRCT; 2.
DR TIGRFAMs; TIGR00574; dnll; 1.
DR PROSITE; PS00172; BRCT; 2.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS00333; DNA_LIGASE_A2; 1.
DR PROSITE; PS00160; DNA_LIGASE_A3; 1.
KW ATP-binding; DNA recombination; DNA repair; DNA replication; Ligase.
SQ SEQUENCE 845 AA; 96487 MW; 9CA8A6B3935C259C CRC64;

```

```

Query Match      39.8%; Score 51; DB 2; Length 845;
Best Local Similarity 47.6%; Pred. No. 3e+02;
Matches 10; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```

```

Oy 4 LGQOPIDRQTIEOMAQLLAEL 24
Db 713 LGIRPSEQQTPEWAPVIADL 733

RESULT 19
ID QBTF7 PRELIMINARY; PRT; 911 AA.
AC QBTF7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:4932416F16 product:DNA LIGASE IV (EC 6.5.1.1)
DE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE).
GN Name-Lig4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).

```


Thu Jan 27 12:40:01 2005

us-09-879-248-6_copy_31_57.rup

Page 10

Db 184 LKQWPLSRKIEALTQITQEWLK 206

Search completed: January 25, 2005, 12:15:27
Job time : 111.769 secs